SPRINGER REFERENCE

Stan Z. Li
Editor

Anil K. Jain
Editorial Advisor

Encyclopedia of Biometrics

Springer

Encyclopedia of Biometrics

Stan Z. Li
Editor
Anil K. Jain
Editorial Advisor

Encyclopedia of Biometrics

With 683 Figures* and 88 Tables

Springer

*For color figures please see our Electronic Reference on www.springerlink.com

Editor
Stan Z. Li
Professor
Center for Biometrics and Security Research
Chinese Academy of Sciences
Beijing
China

Editorial Advisor
Anil Jain
Professor
Department of Computer Science & Engineering
Michigan State University
East Lansing, MI
USA

Library of Congress Control Number: 2009929415

ISBN: 978-0-387-73002-8

This publication is available also as:
Electronic version under ISBN 978-0-387-73003-5
Print and electronic bundle under ISBN 978-0-387-73004-2

© 2009 Springer Science+Business Media, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC., 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.
The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

springer.com

Printed on acid-free paper SPIN: 11878742 — 5 4 3 2 1 0

Preface

The *Encyclopedia of Biometrics* provides a comprehensive reference to concepts, technologies, issues, and trends in the field of biometrics. The volume covers all important aspects – research, development, and applications, including biometric sensors and devices, methods and algorithms, sample quality, system design and implementation, databases, performance testing, information security, and standardization. Leading experts around the world contributed to this collection of over 200 in-depth essays accompanied by more than 600 definitional entries.

The focus of the encyclopedia is on immediate, yet comprehensive, information in an easy-to-use format which is accessible to researchers and scientists, system designers, engineers, programmers, students, practitioners, and government agents working in the broad field of biometrics. It is available as a print edition as well as a fully searchable version with extensive cross-referencing and updates as new trends and terms arise.

Key Features at a Glance

- Serves as an immediate point of entry into the field for in-depth research
- Covers biometrics of face, fingerprints, iris, vein, voice, hand, ear, gait, skin, tongue, dental, odor, skull, and DNA
- A–Z format allows intuitive and easy-to-use access
- Cross-referenced entries
- Internationally renowned editorial board, from across the scientific and engineering disciplines and geographies
- Balanced coverage

Acknowledgments

I am grateful to all the people who have played a part in the production of this encyclopedia. First of all, my deep gratitude to Anil K. Jain, Editorial Advisor for sharing his knowledge and expertise, and important advice and suggestions for this encyclopedia. Our stellar team of area editors has done excellent work in writing, inviting, and reviewing the contributions from many leaders of the field. My sincere thanks to Andy Adler, Joseph Campbell, Raffaele Cappelli, Christophe Champod, Stephen Elliott, Julian Fierrez, Jean-Christophe Fondeur, Carmen Garcia-Mateo, Josef Kittler, Hale Kim, Ajay Kumar, Davide Maltoni, Aleix Martinez, Mark Nixon, Geppy Parziale, Fernando Podio, Salil Prabhakar, Arun Ross, Marios Savvides, Yoichi Seto, Colin Soutar, Wei-Yun Yau, Pong C. Yuen, and David Zhang. I would like to thank these area editors for their help in creating this book and the numerous authors for their individual contributions. Special thanks are due to people at Springer for their enthusiasm, advice, and support. Jennifer Evans, Susan Lagerstrom-Fife, Michaela Bilic, Tina Shelton, and Anil Chandy have played key roles at different times in the development of this book.

Stan Z. Li
Editor-in-Chief

© 2009 Springer Science+Business Media, LLC

Editor

Stan Z. Li
Professor
Center for Biometrics and Security Research
Chinese Academy of Sciences
Beijing
China
szli@cbsr.ia.ac.cn

Editorial Advisor

Anil K. Jain
Professor
Department of Computer Science & Engineering
Michigan State University
East Lansing, MI
USA
jain@cse.msu.edu

Area Editors

Andy Adler
Systems and Computer Engineering
Carleton University
1125 Colonel By Drive
Ottawa, ON K1S 5B6
Canada
adler@sce.carleton.ca

Joseph P. Campbell
Information Systems Technology Group
Massachusetts Institute of Technology
244 Wood Street
Lexington, MA 02420-9108
USA
jpc@ll.mit.edu

Raffaele Cappelli
Department of Electronics, Informatics and Systems (DEIS)
University of Bologna
Via Gaspare Finali
56 Cesena
Italy
cappelli@csr.unibo.it

Christophe Champod
Institut de Police Scientifique
Ecole des Sciences Criminelles
Universite' de Lausanne
1015 Lausanne
Switzerland
Christophe.Champod@unil.ch

Stephen Elliott
Department of Industrial Technology
Purdue University
West Lafayette, IN 47907-2021
USA
elliott@purdue.edu

Julian Fierrez
Escuela Politecnica Superior
Universidad Autonoma de Madrid
28049 Madrid
Spain
julian.fierrez@uam.es

Jean-Christophe Fondeur
Sagem Sécurité
Le Ponant de Paris - 27, rue Leblanc - 75512
Paris Cedex 15
France
jean-christophe.fondeur@sagem.com

Carmen Garcia-Mateo
Telecommunication Engineering School
University of Vigo
36310 Vigo
Spain
carmen@gts.tsc.uvigo.es

Josef Kittler
Faculty of Engineering & Physical Sciences
University of Surrey
Guildford, GU2 7XH
UK
J.Kittler@surrey.ac.uk

Hale Kim
School of Information & Communication Engineering
253 Yonghyun-dong
Nam-ku, Incheon, 402-751
Korea
hikim@inha.ac.kr

Ajay Kumar
Department of Chemical Engineering & Chemistry
The Hong Kong Polytechnic University
Hung Hom, Kowloon
Hong Kong
ajaykr@ieee.org

Davide Maltoni
Department of Electronics, Informatics and Systems (DEIS)
University of Bologna
Viale Risorgimento, 2 Bologna
Cesena
Italy
maltoni@csr.unibo.it

© 2009 Springer Science+Business Media, LLC

Aleix Martinez
Department of Electrical and Computer Engineering
The Ohio State University
2015 Neil Avenue
Columbus, OH 43210
USA
aleix@ece.osu.edu

Mark Nixon
Department of Electronics and Computer Science
University of Southampton
SO17 1BJ
UK
msn@ecs.soton.ac.uk

Geppy Parziale
iNVASIVE CODE
Avda Passapera 17 - 17310 Lloret de Mar (Gl)
Barcelona
Spain
geppy.parziale@invasivecode.com

Fernando Podio
Computer Security Division
National Institute of Standards and Technology (NIST)
Gaithersburg, MD
USA
fernando@nist.gov

Salil Prabhakar
Algorithms Research Group
DigitalPersona
Redwood City, CA 94063
USA
SalilP@digitalpersona.com

Arun Ross
Lane Department of Computer Science and Electrical Engineering
West Virginia University
Morgantown, 26506 WV
USA
Arun.Ross@mail.wvu.edu

Marios Savvides
Department of Electrical and Computer Engineering
Carnegie Mellon University
4720 Forbes Avenue
Pittsburg, PA 15213
USA
msavvid@cs.cmu.edu

Yoichi Seto
Advanced Institute of Industrial Technology
Tokyo Metropolitan University
Tokyo
Japan
seto.yoichi@aiit.ac.jp

Colin Soutar
273, Broadway Avenue
Toronto, ON M4P 1W1
Canada
colin.soutar@rogers.com

Wei-Yun Yau
Agency for Science, Technology & Research
Institute for Infocomm Research
1 Fusionopolis Way
#21-01 Connexis (South Tower)
Singapore 138632
Singapore
wyyau@i2r.a-star.edu.sg

Pong C. Yuen
Department of Computer Science
Hong Kong Baptist University
Kowloon Tong
Hong Kong
pcyuen@comp.hkbu.edu.hk

David Zhang
Department of Computing
The Hong Kong Polytechnic University
Hung Hom, Kowloon
Hong Kong
csdzhang@comp.polyu.edu.hk

List of Contributors

Mongi Abidi
The University of Tennessee
Knoxville, TN 37996–2100
USA
abidi@utk.edu

Mathias M. Adankon
University of Quebec
1100, Notre-Dame Street West
Montreal
Canada
mathias@livia.etsmtl.ca

Andy Adler
Carleton University
1125 Colonel By Drive
K1S 5B6
Ottawa, ON
Canada
adler@sce.carleton.ca

Gaurav Aggarwal
University of Maryland
College Park, MD 20742
USA
gaurav@cfar.umd.edu

Jonathan Agre
Fujitsu Laboratories of America
5801 North Sheridan Road, Suite 19A
College Park, MD 20740
USA
jonathan.agre@us.fujitsu.com

Thomas Albrecht
Computer Science Department
University of Basel
Bernoullistrasse 16, 4056 Basel
Switzerland
thomas.albrecht@unibas.ch

Petar S. Aleksic
Google Inc.
76 9th Avenue
New York, NY 10011
USA
apetar@google.com

Nigel M. Allinson
University of Sheffield
Mappin Street
Sheffield, S1 3JD
UK
n.allinson@sheffield.ac.uk

Fernando Alonso-Fernandez
Biometric Recognition Group - ATVS
Escuela Politecnica Superior
Universidad Autonoma de Madrid
Campus de Cantoblanco
Madrid
Spain
fernando.alonso@uam.es

Nick Bartlow
West Virginia University
Morgantown, WV 26506
USA
nick.bartlow@mail.wvu.edu

Alex Bazin
Fujitsu Services
London, W1U 3BW
UK
alex.bazin@uk.fujitsu.com

George Bebis
University of Nevada
Reno, NV 89557
USA
bebis@cse.unr.edu

Herman Bergman
Certified Fingerprint Expert
San Francisco, USA
hermanbergman@comcast.net

Bir Bhanu
Center for Research in Intelligent Systems
University of California
Riverside, CA 92506
USA
bhanu@cris.ucr.edu

© 2009 Springer Science+Business Media, LLC

Josef Bigun
Halmstad University
IDE, SE–30118
Halmstad
Sweden
josef.bigun@hh.se

Vance Bjorn
DigitalPersona Inc.
720 Bay Road
Redwood City, CA 94063
USA
vanceb@digitalpersona.com

Jean-Francois, Bonastre
Amtel-France
University of Avignon
339 chemin des Meinajariès
BP1228
Cedex 9, Avignon
France
jean-francois.bonastre@univ-avignon.fr

Jeffrey E. Boyd
University of Calgary
Calgary, AB T2N 1N4
Canada
boyd@cpsc.ucalgary.ca

Michael C. Bromby
Glasgow Caledonian University
Cowcaddens Road
G4 0BA Glasgow
UK
M.Bromby@gcal.ac.uk

Randy P. Broussard
United States Naval Academy
121 Blake Rd Annapolis
MD 21402-5000
USA
broussar@usna.edu

Kelvin Bryant
North Carolina A&T State University
1601 East Market Street
Greensboro, North Carolina 27411
USA
ksbryant@ncat.edu

Taihei Munemoto
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3890
USA

Douglas J. Buettner
The Aerospace Corporation
P.O. Box 92957
Los Angeles, CA
USA
Douglas.J.Buettner@aero.org

Anne M. Burrows
Duquesne University
Pittsburgh, PA 15282
USA
burrows@duq.edu

Christoph Busch
Fraunhofer Institute Graphische
Datenverarbeitung
Fraunhoferstraße 5D-64283
Darmstadt
Germany
christoph.busch@igd.fraunhofer.de

Thomas A. Busey
Psychological and Brain Sciences
Program in Cognitive Science
Indiana University
Bloomington Indiana, IN 47405
USA
busey@indiana.edu

James L. Cambier
Cross Match Technologies
3960 RCA Boulevard
Palm Beach Gardens, FL 33410
USA
James.Cambier@crossmatch.com

John W. M. Campbell
Bion Biometrics Inc.
8283 Greensboro Drive, H8056
Ottawa, ON
Canada
john@bionbiometrics.com

Patrizio Campisi
Applied Electronics Department
University of Roma TRE
via della Vasca Navale, 84
Rome 00146
Italy
campisi@uniroma3.it

Greg Cannon
Cross Match Technologies
3960 RCA Boulevard
Palm Beach Gardens, FL 33410
USA
Greg.Cannon@CrossMatch.com

Kai Cao
Institute of Automation
Chinese Academy of Sciences
95 Zhingguancun Donglu
Beijing 100 190
People's Republic of China
caokai@fingerpass.net.cn

Raffaele Cappelli
Department of Electronics, Informatics
and Systems (DEIS)
University of Bologna
Via Gaspare Finali 56 Cesena
Bologna
Italy
cappelli@csr.unibo.it

Carlos D. Castillo
Department of Computer Science
University of Maryland
College Park, MD 20742
USA
carlos@cs.umd.edu

Ann Cavoukian
Office of the Information and Privacy Commissioner
2 Bloor St. E., Suite 1400
Toronto, ON M4W 1A8
Canada
infoipc@ipc.on.ca

Christophe Champod
Institut de Police Scientifique
Ecole des Sciences Criminelles
Université de Lausanne
1015 Lausanne
Switzerland
Christophe.Champod@unil.ch

Rama Chellappa
University of Maryland
College Park, MD 20742-3275
USA
Rama@umiacs.umd.edu

Hong Chen
Michigan State University
East Lansing, MI 48824
USA
chenhon2 @cse.msu.edu

Hui Chen
Center for Research in Intelligent Systems
University of California
Riverside, CA 92506
USA
hchen@vislab.ucr.edu

Yi Chen
Michigan State University
Apt 2B, 3200 Trappers Cove Trail
Lansing, MI 48910
USA
chenyi1@msu.edu

Mohamed Cheriet
Department of Automated Manufacturing
Engineering
École de technologie supérieure
Montréal, Québec H3C 1K3
Canada
Mohamed.cheriet@etsmtl.ca

Girija Chetty
National Centre for Biometric Studies
University of Canberra
ACT 2601
Australia
Girija.Chetty@canberra.edu.au

Alex Hwansoo Choi
Department of Information Engineering
Myongji University
Seoul 137-060
South Korea
alexchoi@tech-sphere.com

Seungjin Choi
Department of Computer Science
Pohang University of Science and Technology
San 31 Hyoja-dong
Nam-gu, Pohang 790-784
Korea
seungjin@postech.ac.kr

Michal Choraś
Institute of Telecommunications
University of Technology and Life Sciences
Bydgoszcz 85-225
Poland
chorasm@atr.bydgoszcz.pl

Jeffrey F. Cohn
University of Pittsburgh
Pittsburgh, PA
USA
jeffcohn@cs.cmu.edu

Raphael Coquoz
Institut De Police Scientifique
Ecole Des Science Criminelles
1015 Laussane
Switzerland
Raphael.Coquoz@unil.ch

Bojan Cukic
West Virginia University
Morgantown, WV 26506
USA
bojan.cukic@mail.wvu.edu

Allison M. Curran
Department of Chemistry and Biochemistry
International Forensic Research Institute
Florida International University
Miami, FL 33199
USA
amcurran@gmail.com

Qionghai Dai
Automation Department
Tsinghua University
Beijing
Pepole's Republic of China
daiqh@tsinghua.edu.cn

Sarat C. Dass
Michigan State University
East Lansing, MI 48824
USA
sdass@msu.edu

John Daugman
University of Cambridge
William Gates Building
15 JJ Thomson Avenue
Cambridge CB3 0FD
UK
John.Daugman@CL.cam.ac.uk

James W. Davis
The Ohio State University
Columbus, OH 43210
USA
jwdavis@cse.ohio-state.edu

David Day
International Biometric Group
New York, NY 100 04
USA
dday@biometricgroup.com

Farzin Deravi
University of Kent
Canterbury, Kent, CT2 7NZ
UK
f.Deravi@kent.ac.uk

Laura Docio-Fernandez
University of Vigo
36310 Vigo
Spain
ldocio@gts.tsc.uvigo.es

Cathryn Downing
University of Cambridge
15 JJ Thomson Avenue
Cambridge CB30FD
UK
cathryn.downing@LL.cam.ac.uk

Gerry Vernon Dozier
College of Engineering
North Carolina A & T State University
1601 E Market Street
Greensboro NC 27411
USA
gvdozier@ncat.edu

Andrzej Drygajlo
Swiss Federal Institute of Technology
Bureau ELE 233
Lausanne CH-101
Switzerland
andrzej.drygajlo@epfl.ch

Nicolae Duta
Nuance Communications
Gold Coast Campus
Burlington, MA 01803
USA
Nicolae.Duta@nuance.com

Ahmed Elgammal
Department of Computer Science
Rutgers University
110 Frelinghuysen Road
Piscataway, NJ 08854
USA
elgammal@cs.rutgers.edu

Stephen J. Elliott
Purdue University
West Lafayette, IN 47907
USA
elliott@purdue.edu

Nicholas W. D. Evans
Swansea University
Singleton Park, Swansea, SA2 8PP
UK
and
Institut Eurécom
2229 route des cretes
Sophia-Antipolis 06 560
France
nicholas.evans@eurecom.fr

Julian Fierrez
Escuela Politecnica Superior
Universidad Autonoma de Madrid
C/ Fco Tomas y Valiente 11
28049 Madrid
Spain
julian.fierrez@uam.es

Jean-Christophe Fondeur
Sagem Sécurité
Le Ponant de Paris - 27
rue Leblanc - 75512, Paris cedex 15
France
jcfondeur@morpho.com

Marc Friedman
Retica Systems Inc.
201 Jones Road, Third Floor West
Waltham, MA 02451
USA
m.friedman@retica.com

Pascal Fua
School of Computer and Communication Science
Ecole Polytechnique Federale de Lausanne
IC-CVLab, Station 14
Lausanne CH-1015
Switzerland
pascal.fua@epfl.ch

Kenneth G. Furton
Department of Chemistry and Biochemistry
International Forensic Research Institute
Florida International University
Miami, FL 33199
USA
furtonk@fiu.edu

Jihyeon Jang
Inha University
253, Yonghyun-dong
Nam-gu Incheon 402-751
Korea
jhjang@vision.inha.ac.kr

Ioannis Patras
Department of Electronic Engineering
Queen Mary University of London
Mile End Rood
London E1 4NS
UK
I.Patras@elec.qmul.ac.uk

Wen Gao
Institute of Computing Technology
Chinese Academy of Sciences
Beijing, People's Republic of China
and
Peking University
Beijing, People's Republic of China
Wgao@pku.edu.cn

Lu Gao
CDM Optics, Inc.
4001 Discovery Dr., Suite 130
Boulder, CO 80303
USA
lu.gao@cdm-optics.com

Carmen Garcia-Mateo
University of Vigo
36310, Vigo
Spain
carmen@gts.tsc.uvigo.es

Sonia Garcia-Salicetti
TELECOM SudParis
9 Rue Charles Fourier
Evry 91011
France
Sonia.Salicetti@int-edu.eu

Xin Geng
Deakin University
Melbourne, VIC 3125
Australia
xgeng @deakin.edu.au

Shaogang Gong
Queen Mary
University of London
Mile End Road
London E1 4NS
UK
sgg@dcs.qmul.ac.uk

Javier Ortega-Garcia
Biometric Recognition Group - ATVS
Escuela Politecnica Superior
Universidad Autonoma de Madrid
Campus de Cantoblanco
Madrid 28049
Spain
Javier.Ortega@uam.es

Joaquín González-Rodríguez
Area de Tratamiento de Voz y Señales (ATVS)
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Madrid 28049
Spain
joaquin.gonzalez@uam.es

Dmitry O. Gorodnichy
Laboratory and Scientific Services
Directorate
Canada Border Services Agency
79 Bently Ave.
Ottawa, ON, K1A OL5
Canada
dg@videorecognition.com

Joseph van der Gracht
Holospex Inc.,
Columbia, MD 21044
USA
vanderj@holospex.com

Patrick Grother
National Institute of Standards and Technology
100 Bureau Drive, MS 8930
Gaithersburg, MD 20899
USA
patrick.grother@nist.gov

Leon Gu
Carnegie Mellon University
Pittsburgh, PA 15217
USA
gu@cs.cmu.edu

David L. Guyton
The Wilmer Ophthalmological Institute
The Johns Hopkins University School of Medicine
Baltimore, MD 21287-9028
USA
dguyton@jhmi.edu

Abdenour Hadid
Machine Vision Group
Department of Electrical and Information Engineering
University of Oulu
P.O. Box 4500
Oulu 90014
Finland
hadid@ee.oulu.fi

Onur C. Hamsici
The Ohio State University
205 Dreese Laboratory
2015 Neil Avenue 43210
Columbus, OH 43210
USA
hamsicio@ece.osu.edu

Seiichiro Hangai
Department of Electrical Engineering
Tokyo University of Science
1-3 Kagurazaka, Shinjyuku-ku
Tokyo 162-8601
Japan
hangai@ee.kagu.sut.ac.jp

Masanori Hara
NEC Corporation
5–7–1 Shiba, Minato-ku
Tokyo 108-8001
Japan
m-hara@da.jp.nec.com

Dominique, Harrington
Tygart Technology Inc.
1543 Fairmont Ave
Fairmont, WV 26554
USA
dc.whitewolf@verizon.net

Jean Hennebert
Department of Informatics, University of Fribourg
1700 Fribourg
Switzerland
jean.hennebert@unifr.ch
and
Institute of Business Information Systems HES-SO Valais, TechnoArk 3, 3960 Sierre
Switzerland
jean.hennebert@hevs.ch

Olaf Henniger
Fraunhofer Institute for Secure Information Technology
Rheinstr. 75, 642957 Darmstadt
Germany
henniger@sit.fraunhofer.de

Javier Hernando
Technical University of Catalonia
Building D5, Campus Nord UPC
Jordi Girona 1–3
Barcelona 08034
Spain
javier@gps.tsc.upc.edu

Fred Herr
NIST Contractor - Identification Technology Partners Inc.
North Potomac, MD 20878
USA
fherr@idtp.com

R. Austin Hicklin
Noblis
3150 Fairview Park Drive
Falls Church, VA 22043
USA
hicklin@noblis.org

Tacha Hicks
Ecole des Sciences Criminelles / Institut de Police Scientifique
Université de Lausanne
Bâtiment Batochime
Lausanne CH-1015
Switzerland
tnhc@mac.com

Peter T. Higgins
Higgins, Hermansen, Banikas
LLC 2111 Wilson Blvd, Suite 607
Arlington, VA 22201
USA
peter.higgins@thehhb.com

Mitsutoshi Himaga
Hitachi-Omron Terminal Solutions, Corp.
1-6-3 Osaki, Shinagawa Ward
Tokyo 141-8576
Japan
mitsutoshi_himaga@hitachi-omron-ts.com

Tom Hopper (retired)
Criminal Justice Information Services Division
FBI
thopper@leo.gov

Nesma Houmani
TELECOM SudParis
9 Rue Charles Fourier
Evry 91011
France
Nesma.Houmani@int-edu.eu

Wen-Hsing Hsu
National Tsing Hua University
Taiwan
People's Republic of China
whhsu@ee.nthu.edu.tw

Dunxu Hu
Department of Computer Science
Rutgers University
Piscataway, NJ, 08902
USA
hu39@cs.purdue.edu

David J. Hurley
School of Electronics and Computer Science
University of Southampton
SO 17 1BJ
UK
djh@AnalyticalEngines.co.uk

Kristina Irsch
The Wilmer Ophthalmological Institute
The Johns Hopkins University
School of Medicine
Baltimore, MD 21287-9028
USA
kirsch1@jhmi.edu

David W. Jacobs
Department of Computer Science
University of Maryland
College Park, MD 20742
USA
djacobs@cs.umd.edu

Anil K. Jain
Department of Computer Science and Engineering
Michigan State University
3115, Engineering Building
East Lansing, MI 48824
USA
jain@cse.msu.edu

Kui Jia
Shenzhen Institute of Advanced Integration
Technology
CAS / CUHK
Shenzhen
People's Republic of China
kui.jia@sub.siat.ac.cn

Xudong Jiang
School of Electrical and Electronic Engineering
Nanyang Technological University
Nanyang Link 639798
Singapore
exdjiang@ntu.edu.sg

Zhong Jin
School of Computer Science and Technology
Nanjing University of Science and Technology
Nanjing 210094
Peoples Republic of China
zhongjin@mail.njust.edu.cn

Juergen Schroeter
AT&T Labs – Research
180 Park Ave., Room D163
Florham Park, NJ 07932–0971
USA
schroeter@att.com

Ioannis A. Kakadiaris
Department of Computer Science
Department of Electrical and Computer Engineering
and
Department of Biomedical Engineering
University of Houston
MS CSC 3010, 4800 Calhoun
Houston, TX 77204-3010
USA
IKakadia@Central.UH.EDU

Nathan Kalka
West Virginia University
Morgantown
WV 26506
USA
nathan.kalka@mail.wvu.edu

Takeo Kanade
Robotics Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15217
USA
Takeo.Kanade@cs.cmu.edu

Vivek Kanhangad
Hong Kong Polytechnic University
Hung Hom, Kowloon
Hong Kong
csvivek@comp.polyu.edu.hk

Mark Keck
The Ohio State University
Columbus, OH 43210
USA
keck@cse.ohio-state.edu

Robert B. Kennedy (retired)
Royal Canadian Mounted Police
Ottawa, ON K1A OR2
Canada
robertkennedy7@rogers.com

Lauren R. Kennell
The John Hopkins University
Applied Physics Laboratory
Mail Stop 24-W211
11100 John Hopkins Rd.
Laurel, MD 20723-6099
USA
Kennel@usna.edu

Hale Kim
Inha University
Incheon, 402–751
Korea
hikim@inha.ac.kr

Jong Kyoung Kim
Department of Computer Science
Pohang University of Science and Technology
San 31 Hyoja-dong
Nam-gu, Pohang 790-784
Korea
blkimjk@postech.ac.kr

Kye-Hyeon Kim
Department of Computer Science
Pohang University of Science and Technology
San 31 Hyoja-dong
Nam-gu, Pohang 790-784
Korea
fenrir@postech.ac.kr

Naohisa Komatsu
Waseda University
Shinjuku-ku, Tokyo 169–8555
Japan
nkomatsu@waseda.jp

Eric P. Kukula
Purdue University
Knoy Hall of Technology, Room 377D
401 N. Grant Street
West Lafayette, IN 47906–2021
USA
kukula@purdue.edu

Ajay Kumar
Department of Computing
The Hong Kong Polytechnic University
Hung Hom
Hong Kong
ajaykr@ieee.org

J. H. Lai
Department of Mathematics
Sun Yat-sen University
P.O. Box 4500
Guangzhou 510275
People's Republic of China
stsljh@sysu.edu.cn

Kin-Man Lam
Department of Electronic and Information Engineering
The Hong Kong Polytechnic University
Hong Kong
People's Republic of China
enkmlam@inet.polyu.edu.hk

John Larmouth
University of Salford
Salford, Greater Manchester
UK
j.larmouth@btinternet.com

Peter K. Larsen
Laboratory of Biological Anthropology
Institute of Forensic Medicine
University of Copenhagen
DK–2200 Copenhagen
Denmark
pklarsen@sund.ku.dk

Richard T. Lazarick
CSC Identity Labs
Fairmont, WV 26554
USA
rlazarick@csc.com

Victor Lee
International Biometric Group
One Battery Park Plaza
New York, NY 10004
USA
vlee@biometricgroup.com

Graham Leedham
School of Information and
Communication Technology
Griffith University
Gold Coast Campus
Queensland 4222
Australia
g.leedham@gmail.com

Zhen Lei
Centre for Biometrics and Security Research &
National Laboratory of Pattern Recognition
Institute of Automation
Chinese Academy of Sciences
Beijing
People's Republic of China
zlei@cbsr.ia.ac.cn

Yung-Hui Li
Language Technology Institute
Carnegie Mellon University
Pittsburgh, PA 15213
USA
yunghui@cmu.edu

Stan Z. Li
Centre for Biometrics and Security Research
Chinese Academy of Sciences
95 Zhongguancun Donglu
Beijing
People's Republic China
szli@cbsr.ia.ac.cn

Peng Li
Institute of Automation
Chinese Academy of Sciences
95 Zhongguancun Donglu
Beijing 100190
China
lipeng@fingerpass.net.cn

Jianjie Li
Institute of Automation
Chinese Academy of Sciences
95 Zhongguancun Donglu
Beijing 100190
China
lijj@fingerpass.net.cn

Zhouchen Lin
Microsoft Research Asia
Beijing
People's Republic of China
zhoulin@microsoft.com

James J. Little
University of British Columbia
Vancouver, BC V6T 1Z4
Canada
little@cs.ubc.ca

Laura L. Liu
Biometrics Research Centre
The Hong Kong Polytechnic University
Kowloon
Hong Kong
csliliu@comp.polyu.edu.hk

Zongyi Liu
Amazon.com
Seattle, WA 98104
USA
lzongyi@yahoo.com

Marcel Luthi
University of Basel
Bernoullistrasse 16, 4056 Basel
Switzerland
marcel.luethi@unibas.ch

Guangming Lu
School of Computer Science and Technology
Shenzhen Graduate School
Harbin Institute of Technology
Shenzhen
People's Republic of China
Luguangm@hit.edu.cn

Niels Lynnerup
Laboratory of Biological Anthropology
Institute of Forensic Medicine
University of Copenhagen
Copenhagen DK-2200
Denmark
nlynnerup@antrolab.ku.dk

Emanuele Maiorana
University of Roma TRE
via della Vasca Navale, 84
Rome 00146
Italy
maiorana@uniroma3.it

Sotiris Malassiotis
Informatics and Telematics Institute
Center for Research and Technology Hellas
Thermi-Panorama Road
Thermi-Thessaloniki, 57001
Greece
malasiot@iti.gr

Davide Maltoni
Department of Electronics, Informatics and Systems (DEIS)
University of Bologna
Viale Risorgimento
2 Bologna, Cesena
Italy
maltoni@csr.unibo.it

Judith Markowitz
J. Markowitz, Consultants
5801 North Sheridan Road
Suite 19A
Chicago, IL 60660
USA
judith@jmarkowitz.com

Alvin F. Martin
National Institute of Standards and Technology
100 Bureau Drive, Stop 8940
Gaithersburg, MD 20899-8940
USA
alvin.martin@nist.gov

Aleix M. Martinez
Department of Electrical and Computer Engineering
The Ohio State University
2015 Neil Avenue
Columbus, OH 43210
USA
aleix@ece.osu.edu

Marcos Martinez-Diaz
Biometric Recognition Group - ATVS
Escuela Politecnica Superior
Universidad Autonoma de Madrid
Campus de Cantoblanco
Madrid 28049
Spain
marcos.martinez@uam.es

John S. D. Mason
Swansea University
Singleton Park
Swansea, SA2 8PP
UK
j.s.d.mason@swansea.ac.uk

James R. Matey
Electrical and Computer Engineering Department
Maury Hall, MS 14B
Annapolis, MD 21402
USA
matey@usna.edu

Driss Matrouf
University of Avignon
LIA - BP1228, 339 Ch. des Meinajari'es
Avignon 84911
France
driss.matrouf@univ-avignon.fr

Takashi Matsumoto
Waseda University
Shinjuku-ku
Tokyo 169–8555
Japan
takashi@mse.waseda.ac.jp

Rene Mciver
389 Keewatin Avenue
Toronto, ON
M4P 2A4
Canada
Rene.mciver@verkis.com

Gerard G. Medioni
Department of Computer Science
University of Southern California
Los Angeles, CA
USA
medioni@iris.usc.edu

Didier Meuwly
Nederlands Forensisch Instituut
PO Box 6109
The Hague
Netherlands
d.meuwly@nfi.minjus.nl

Krystian Mikolajczyk
Faculty of Engineering & Physical Sciences
University of Surrey
3115 Engineering Building
Guildford
UK
k.mikolajczyk@surrey.ac.uk

Feng Min
Institute for Pattern Recognition and Artificial Intelligence
Huazhong University of Science and Technology
1037 Luoyu Road, Wuhan
People's Republic of China
and
Lotus Hill Institute for Computer Vision and Information Science
People's Republic of China
fmin.lhi@gmail.com

Philippos Mordohai
University of Pennsylvania
Philadelphia, PA 19146
USA
mordohai@seas.upenn.edu

Ken Moses
Forensic Identification Services
130 Hernandez Avenue
San Francisco, CA 94127
USA
forensicid@sbcglobal.net

Crystal Muang
Department of Computer Science
Rutgers University
Piscataway, NJ 08902
USA

Robert Mueller
Giesecke & Devrient GmbH
Prinzregentenstr. 159
Muenchen 81677
Germany
robert.mueller@gi-de.com

Tanvir Singh Mundra
Biometrics Research Laboratory
Department of Electrical Engineering
Indian Institute of Technology
New Delhi
India
tanvirmundra@gmail.com

Manabu Nakano
Information-technology Promotion Agency (IPA)
Bunkyo-ku
Japan
Tokyo 113–6591
mn-naka@ipa.go.jp

Karthik Nandakumar
Institute for Infocomm Research, A*STAR
Fusionopolis
Singapore
knandakumar@izr.a-star.edu.sg

Ramkumar Narayanswamy
Omni Vision CDM Optics, Inc.
4001 Discovery Dr. Boulder
CO 80303
USA
ramkumarn@cdm-optics.com

Alessandro Neri
University of Roma TRE
via della Vasca Navale, 84
Rome 00146
Italy
neri@uniroma3.it

Mark S. Nixon
School of Electronics and Computer Science
University of Southampton
SO17 1BJ
UK
msn@ecs.soton.ac.uk

Javier Ortega-Garcia
Biometric Recognition Group - ATVS, Escuela Politecnica Superior
Universidad Autonoma de Madrid
Madrid
Spain
Javier.Ortega@uam.es

Lisa Osadciw
Department of Electrical Engineering and Computer Science
Syracuse University
Syracuse, NY
USA
laosadci@syr.edu

Hisao Ogata
Hitachi-Omron Terminal Solutions
Corp. Owari-asahi City Aichi 488-8501
Japan
hisao_ogata@hitachi-omran-ts.co

Chen Tai Pang
Institute for Infocomm Research, A*STAR
21 Heng Mui Keng Terrace, 119613 Singapore
Singapore
tpchen@i2r.a-star.edu.sg

Pankanti Sharat
IBM T.J. Watson Research Center
Hawthorne, NY 10532
USA
sharat@us.ibm.com

Maja Pantic
Imperial College London
Department of Computing
London SW7 2AZ
UK
m.pantic@imperial.ac.uk

Sung W. Park
Carnegie Mellon University
Pittsburgh, PA 15213
USA
sungwonp@cmu.edu

Geppy Parziale
iNVASIVE CODE
Avda Passapera 17-H310 Lloretde Mar (GI)
Barcelona, Spain
geppy.parziale@invasivecode.com

Georgios Passalis
University of Houston
Houston, TX 77204
USA
passalis@di.uoa.gr

V. Paúl Pauca
Department of Computer Science
Wake Forest University
4001 Discovery Dr. Suite 130
Winston-Salem, NC
USA
paucavp@wfu.edu

Nikola Pavesic
Faculty of Electrical Engineering
University of Ljubljana
Trzaska 25
SI-1000 Ljubljana
Slovenia
nikola.pavesic@fe.uni-lj.si

Maria Pavlou
Department of Electronic and Electrical Engineering
University of Sheffield
Mappin Street, Sheffield
UK
m.pavlou@sheffield.ac.uk

Xiaoming Peng
College of Automation
University of Electronic Science and Technology of China
Chengdu, Sichuan Province
People's Republic of China
pengxm@uestc.edu.cn

Takis Perakis
Department of Electrical and Computer Engineering and
Department of Biomedical Engineering
University of Houston
Houston, TX 77204
USA
takis@antinoos.gr

Matti Pietikäinen
Department of Electrical and Information Engineering
University of Oulu
Oulu
Finland
mkp@ee.oulu.fi

Fernando Podio
Security Technology Group, Computer Security Division
National Institute of Standards and Technology
12208 Pueblo Road
Gaithersburg, MD
USA
fernando.podio@nist.gov

Norman Poh
CVSSP, FEPS University of Surrey Guilford
Surrey, GU2 7XH
UK
normanpoh@ieee.org

Frank Pollick
Department of Psychology
University of Glasgow
58 Hillhead Street
Glasgow
UK
frank@psy.gla.ac.uk

Prabhakar S.
Algorithms Research Group
DigitalPersona
Redwood City, CA 94063
USA
SalilP@digitalpersona.com

Ryan Rakvic
Department of Electrical Engineering
United States Naval Academy
105 Maryland Avenue
Annapolis, MD
USA
rakvic@usna.edu

Narayanan Ramanathan
University of Maryland
College Park, MD 20742
USA
ramanath@umiacs.umd.edu

Daniel Ramos
ATVS Universidad Autónoma de Madrid
Ciudad Universitaria de Cantoblanco
Madrid
Spain
Daniel.ramos@uam.es

Marek Rejman-Greene
Biometrics Centre of Expertise
Home Office Scientific Research Branch
St Albans, Hertfordshire AL4 9HQ
UK
marek.rejman-greene@homeoffice.gsi.gov.uk

Douglas Reynolds
Information Systems Technology Group
Massachusetts Institute of Technology
244 Wood Street
Lexington, MA
USA
dar@ll.mit.edu

Damon L. Woodard
School of Computing
Clemson University
Clemson, SC 29634-0974
USA
woodard@clemson.edu

Karl Ricanek
Department of Computer Science
University of North Carolina at Wilmington
Wilmington, North Carolina
USA
ricanekk@uncw.edu

Ruben Vera Rodriguez
Swansea University
Singleton Park, Swansea, SA2 8PP
UK
r.vera-rodriguez.405831@swansea.ac.uk

Fabio Roli
Department of Electrical and Electronic Engineering
University of Cagliari
Pizza d'Armi, I 09123 Cagliari
Italy
roli@diee.unica.it

Arun Ross
Lane Department of Computer Science and Electrical Engineering
West Virginia University
Morgantown, 26506 WV
USA
Arun.Ross@mail.wvu.edu

Besma Roui-Abidi
The University of Tennessee
Knoxville, TN 37996–2100
USA
besma@utk.edu

Amit K. Roy-Chowdhury
Department of Electrical Engineering
University of California
Riverside, CA
USA
amitrc@ee.ucr.edu

Anthony Russo
Atrua Technologies, Inc.
1696 Dell Avenue
Campbell, CA 95008
USA
arusso@atrua.com

Raul Sanchez-Reillo
Electronics Technology
University Carlos III of Madrid
Avda. Universidad, 3028911
Leganes, Madrid
Spain
rsreillo@ing.uc3m.es

Aswin Sankaranarayanan
Center for Automation Research
Department of Electrical and Computer Engineering
University of Maryland
College Park, MD 20742
USA
saswin@gmail.com

Sudeep Sarkar
Computer Science and Engineering
University of South Florida
Tampa, FL
USA
sarkar@cse.usf.edu

Marios Savvides
Electrical and Computer Engineering
Carnegie Mellon University
737 Engineering Sciences Building
Pittsburgh, PA 15213-3890
USA
marioss@andrew.cmu.edu

Natalia A. Schmid
West Virginia University
Morgantown, WV 26506
USA
Natalia.Schmid@mail.wvu.edu

Bethany Schneider
Psychology Department
Indiana University
P.O. Box 530
Bloomington, IN 47405
USA
bschneid@indiana.edu

Adee A. Schoon
Animal Behavior Group
Leiden University
P.O. Box 9516, 2300RA Leiden
The Netherlands
adee.schoon@klpd.politie.nl

Michael Schuckers
Department of Mathematics, Computer Science and Statistics
St. Lawrence University
401 N. Grant Street
Canton, NY
USA
schuckers@stlawu.edu

Stephanie Schuckers
Department of Electrical and Computer Engineering
Clarkson University
PO Box 5720
Potsdam, NY
USA
sschucke@clarkson.edu

Dale Setlak
AuthenTec, Inc.
PO Box 6109
Melbourne, FL
USA
dsetlak@authentec.com

Yoichi Seto
Laboratory 162-a
Advanced Institute of Industrial Technology
Kita 14, Nishi 9, Kita-ku
Tokyo
Japan
seto.yoichi@ailt.ac.jp

Shiguang Shan
Institute of Computing Technology
Chinese Academy of Sciences
Beijing
People's Republic of China
sgshan@ict.ac.in

Vinay Shama
The Ohio State University
Columbus, OH 43210
USA
sharmav@cse.ohio-state.edu

Koichi Shimizu
Bioengineering and Bioinformatics
Graduate School of Information S&T
Hokkaido University
Kita 14, Nishi 9, Kita-ku
Sapporo 060-0814
Japan
shimizu@bme.ist.hokudai.ac.jp

Paulo E. X. Silveira
CDM Optics, Inc.
4001 Discovery Dr. Suite 130
Boulder, CO 80303
USA
paulos@cdm-optics.com

Terence Sim
School of Computing
National University of Singapore
East Road 95#, Haidian Dist., Zhongguancun
Singapore
tsim@comp.nus.edu.sg

Richa Singh
West Virginia University
Morgantown, WV 26506
USA
richas@csee.wvu.edu

Kelly Smith
West Virginia University
Morgantown, WV 26506
USA
smith@biometrics.org

Kate Smith-Miles
Deakin University
Melbourne, VIC 3125
Australia
katesm@deakin.edu.au

Jung Soh
Sun Center of Excellence for Visual Genomics
University of Calgary
Calgary, Alberta
Canada
jsoh@ucalgary.ca

Colin Soutar
273, Broadway Avenue
Toronto, ON M4P 1W1
Canada
colin.soutar@rogers.com

Philip Statham
Biometrics Consultant
1 Blueberry Road
Charlton Kings Cheltenham
UK
philip.statham@talktalk.net

Alex Stoianov
Office of the Information and Privacy Commissioner
2 Bloor St. E., Suite 1400
Toronto, ON M4W 1A8
Canada
alexchoi@tech-sphere.com

Jinli Suo
Lotus Hill Institute for Computer Vision and Information Science
City of Ezhou, Hubei Province
Public Republic of China
jlsuo.lhi@gmail.com

Matthew Swayze
Daon, Inc.
11955 Freedom Dr., Suite 16000
Reston, VA
USA
matthew.swayze@daon.com

Elham Tabassi
Information Access Division
National Institute of Standards and Technology
100 Bureau Drive Stop 1020
Gaithersburg, MD
USA
elham.tabassi@nist.gov

Samir Tamer
Ingersoll Rand Recognition Systems Inc.
1520 Dell Avenue
Campbell, CA
USA
Samir_Tamer@irco.com

Xunqiang Tao
Centre for Biometrics and Security Research
The key Laboratory of Complex System and Intelligence Science
Chinese Academy of Sciences
Institute of Automation
95 Zhingguancun Denglu, Beijing 100190
People's Republic of China
taoxunqiang@fingerpass.net.cn

Chin-Hung Teng
Yuan Ze University
Taiwan
People's Republic of China
chteng@saturn.yzu.edu.tw

Andrew Beng Jin Teoh
Biometrics Engineering Research Center (BERC)
School of Electrical and Electronic Engineering
Yonsei University
Seoul
South Korea
bjteoh@yonsei.ac.kr

Theoharis Theoharis
Department of Computer Science
Department of Electrical and Computer Engineering
and
Department of Biomedical Engineering, University of Houston
Houston, TX 77204
USA
theotheo@di.uoa.gr

Michael Thieme
International Biometric Group
New York, NY 10004
USA
mthieme@biometricgroup.com

Jason Thornton
Electrical & Computer Engineering
Carnegie Mellon University
19 William Road
Pittsburgh, PA 15213
USA
jthornto@gmail.com

Jie Tian
Center for Biometrics and Security Research Science
Chinese Academy of Sciences
Institute of Automation, 95
Zhingguancun Donglu
Beijing 100190
People's Republic of China
tian@ieee.org

Catherine J. Tilton
Daon, Inc.
11955 Freedom Drive, Suite 16000
Reston, VA
USA
Cathy.Tilton@daon.com

Massimo Tistarelli
Computer Vision Laboratory, Facoltà di Architettura di Alghero, Dipartimento di
Università di Sassari
palazzo del Pou Salit - Piazza Duomo 6
Alghero (SS)
Italia
tista@uniss.it

George Toderici
Department of Computer Science
Department of ECE and Department of Biomedical Engineering
University of Houston
Houston, TX 77204
USA
Toderici@cs.uh.edu

Doroteo T. Toledano
Universidad Autonoma de Madrid (UAM)
Ciudad Universitaria de Cantoblanco
Madrid
Spain
doroteo.torre@uam.es

Ted Tomonaga
Konica Minolta Technology Center, Inc.
Tokyo 191–8511
Japan
ted.t@konicaminolta.jp

Todd C. Torgersen
Wake Forest University
Winston-Salem, NC 27109
USA
torgerse@wfu.edu

Yasunari Tosa
Retica Systems, Inc.
Waltham, MA 02451
USA
ytosa@retica.com

Bori Toth
Deloitte & Touche LLP
London, EC4A ATR
UK
aboritoth@gmail.com

Alessandro Triglia
OSS Nokalva, Inc.
22 Baker St
Somerset, NJ 08873
USA
sandro@oss.com

Ryan Triplett
Biometric Services International, LLC
A National Biometric Security Project Company
150 Clay Street, Suite. 350
Morgantown, WV
USA
rtriplett@nationalbiometric.org

Vitomir Štruc
Faculty of Electrical Engineering
University of Ljubljana, Tržaška 25
SI–1000, Ljubljana
Slovenia
vitomir.struc@fe.uni-lj.si

Tinne Tuytelaars
Department of Electrical Engineering
Katholieke Universiteit Leuven
Kasteelpark Arenberg 10
B–3001 Leuven
Belgium
Tinne.Tuytelaars@esat.kuleuven.be

Ambrish Tyagi
The Ohio State University
Columbus, OH 43210
USA
tyagia@cse.ohio-state.edu

Kaoru Uchida
Mobile Terminals Technologies Division
NEC Corporation
211 Valentine Hall
Kawasaki
Japan
k-uchida@bc.jp.nec.com

David Usher
Retica Systems Inc.
201 Jones Road Third Floor West Waltham
MA 02451
USA
dusher@retica.com

Mayank Vatsa
West Virginia University
Morgantown, WV 26506
USA
mayankv@csee.wvu.edu

Kalyan Veeramachaneni
Syracuse University
Syracuse, NY 13244
USA
kveerama@syr.edu

Ashok Veeraraghavan
University of Maryland
College Park, MD 20742
USA
vashok@umiacs.umd.edu

Thomas Vetter
Departement Informatik
Universitaet Basel
Bernoullistr. 16
Basel, 4056
Switzerland
thomas.vetter@unibas.ch

B. V. K. Vijaya Kumar
Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
USA
kumar@ece.cmu.edu

Michael Wagner
School of Information Sciences
University of Canberra, ACT, 2601
Australia
michael.wagner @canberra.edu.au

Liang Wan
Department of Electronic Engineering
City University of Hong Kong
Kowloon
Hong Kong
lwan@cse.cuhk.edu.hk

Masaki Watanabe
Fujitsu Laboratories Ltd.
1–10–40 Higashi-oi
Shinagawa-ku
Kawasaki
Japan
m.watanabe@jp.fujitsu.com

James L. Wayman
College of Engineering
San Jose State University
San Jose, CA 95192-0205
USA
James.Wayman@sjsu.edu

Lior Wolf
The Blavatnik School of Computer Science
Tel-Aviv
Ramat Aviv, Tel Aviv 69978
Israel
wolf@cs.tau.ac.il

Damon L. Woodard
School of Computing
Clemson University
Clemson, SC
USA
woodard@clemson.edu

Xudong Xie
Automation Department
Tsinghua University
Beijing
People's Republic of China
xdxie@mail.tsinghua.edu.cn

Yilei Xu
Department of Electrical Engineering
University of California
Riverside, CA
USA
yxu@ee.ucr.edu

Asahiko Yamada
Toshiba Solutions Corporation
9 Belland Drive
Tokyo
Japan
Yamada.Asahiko@toshiba-sol.co.jp

Brian A. Yamashita
Forensic Identification Operations Support Services
National Services and Research
Royal Canadian Mounted Police
Ottawa, ON
Canada
brian.yamashita@rcmp-grc.gc.ca

Chew-Yean Yam
University of Southampton
Southampton, SO17 1BJ
UK
cyy@ecs.soton.ac.uk

Weilong Yang
Center for Biometrics and Security Research &
National Laboratory of Pattern Recognition
Institute of Automation
Chinese Academy of Sciences
Beijing
People's Republic of China
wlyang@cbsr.ia.ac.cn

Jian Yang
School of Computer Science and Technology
Nanjing University of Science and Technology
Nanjing
People's Republic of China
csjyang@mail.njust.edu.cn

Jingyu Yang
School of Computer Science and Technology
Nanjing University of Science and Technology
Nanjing, 210094
People's Republic of China
yangjy@mail.njust.edu.cn

Xin Yang
Center for Biometrics and Security Research
Chinese Academy of Sciences
Institute of Automation
95 Zhingguancun Donglu
Beijing 100190
People's Republic of China
xin.yang@ia.ac.cn

Ming-Hsuan Yang
Honda Research Institute
800 California Street
Mountain View, CA
USA
mhyang@ieee.org

Wei-Yun Yau
Institute for Infocomm Research
Agency for Science, Technology & Research
Singapore 119613
Singapore
wyyau@i2r.a-star.edu.sg

Dong Yi
Biometrics and Security Research &
National Laboratory of Pattern Recognition
Institute of Automation
Chinese Academy of Sciences
Beijing
People's Republic of China
dyi@cbsr.ia.ac.cn

Isao Yoshimura
Tokyo University of Science
Shinjuku-ku, Tokyo 162–8501
Japan
isao.yoshimura@m3.dion.ne.jp

Mitsu Yoshimura
Ritsumeikan University
Sakyo-ku, Kyoto 603–8577
Japan
yosimuramt@k9.dion.ne.jp

Pong C. Yuen
Hong Kong Baptist University
Kowloon Tong
Hong Kong
pcyuen@comp.hkbu.edu.hk

Arie Zeelenberg
Fingerprint Department
National Police Force, Natuursteenlaan 73
27 19TB Zoetermeer
The Netherlands
wasarie@wanadoo.nl

David Zhang
Department of Computing
The Hong Kong Polytechnic University
Hung Hom, Kowloon
Hong Kong
csdzhang@comp.polyu.edu.hk

Yangyang Zhang
Center for Biometrics and Security Research & The Key Laboratory of Complex System and Intelligence Science, Chinese Academy of Sciences
Institute of Automation
95 Zhingguancun Donglu, Beijing 100190
People's Republic of China
zhangyy@fingerpass.net.cn

Gregory Zektser
U.S. Department of Defence
Biometrics Task Force, 8283 Greensboro Drive, H8056
McLean, VA 22102
USA
zektser_gregory@bah.com

Wei-Shi Zheng
Sun Yat-sen University
Guangzhou, 510275
People's Republic of China
wszheng@ieee.org

Sujing Zhou
Center for Biometrics and Security Research & The Key Laboratory of Complex System and Intelligence Science, Chinese Academy of Sciences
Institute of Automation
95 Zhingguancun Donglu, Beijing 100190
People's Republic of China
Xin.yang@ia.ac.cn

Mingquan Zhou
Institute of Information Science and Technology
Beijing Normal University
#19 Xinjiekouwai Street, Beijing
People's Republic of China
mqzhou@bnu.edu.cn

S. Kevin Zhao
Siemens Corporate Research,
Princeton, NJ 08540
USA
kzhou@scr.siemens.com

Zhi-Hua Zhou
National Key Laboratory for Novel Software Technology
Nanjing University
Nanjing 210093
People's Republic of China
zhouzh@nju.edu.cn

Song-Chun Zhu
Lotus Hill Institute for Computer Vision and Information Science
People's Republic of China
and
Department of Statistics and Department of Computer Science
University of California
Los Angeles, CA
USA
sczhu@stat.ucla.edu

Zhu Yengfang
Discover Financial Services
Riverwoods, IL 60015
USA
yongfangzhu@discover.com

Jasenko Zivanov
University of Basel
4056 Basel
Switzerland
jasenko.zivanov@stud.unibas.ch

A

Abstract Syntax Notation One

A notation commonly used to define abstract syntax and semantics of the data structures (formats), to convey messages in computer communication (similar to XSD, but using a simpler syntax and allowing binary as well as XML encodings of the data). The definitions are independent of any programming language, but it is common for tools to map them into specific programming languages such as C, C++ and Java, and to provide run-time modules that will allow values of those data structures to be encoded in a standardized form and sent to other computer systems. Once defined at the abstract level, ASN.1 messages can be conveyed (through standards and tool support) into a variety of encodings, including very compact binary encodings and human-readable XML encodings.

► Biometric Technical Interface, Standardization

ACBio instance

An ACBio instance is a data structure generated by a Biometric Processing Unit (BPU) and contains data that can be used by an external "validator" process of a biometric verification process to authenticate the BPU and its functional transaction data and other information. The validating data is typically in the form of digitally signed certificates that authenticate the BPU entity and, where appropriate, security relevant aspects of its performance. Examples include: authentication of stored biometric references (templates) used for verification, and certification of the performance capabilities of the biometric technology used in the recognition process. The ACBio instance is digitally signed and bound to the functional transaction data from the BPU. The digitally signed certificates are provided by a trusted 3rd party organisation through an evaluation and certification procedure that is not defined by the ACBio standard.

A BPU is a functional component of a biometric transaction system that operates at a uniform security level. It contains one or more subprocesses, the last of which generates the ACBio instance for the BPU and outputs it together with the functional transaction data from the BPU. The ACBio instance can be used by a subsequent "validator" process of a biometric verification process to authenticate the BPU and its functional transaction data and other information.

► Biometric Security, Standardization

Acceleration

Acceleration of pen-movement during the signing process.

This feature is used for on-line signature verification. There are two ways to obtain the acceleration feature. One way is to measure the pen acceleration directly with accelerometers integrated into the pen. The other way is to compute the acceleration from other measurements, for example, from the second-order derivative of the pen position signal with respect to time.

► Signature Recognition

Access Control

► Access Control, Logical
► Access Control, Physical

© 2009 Springer Science+Business Media, LLC

Access Control, Logical

Vance Bjorn
DigitalPersona Inc., Redwood City, CA, USA

Synonym

Logon, Password management

Definition

Logical access control is the means and procedures to protect access to information on PCs, networks, and mobile phones. A variety of credential types may be used, such as passwords, tokens, or biometrics, to authenticate the user. These credentials may represent something the user knows (password), something the user has (token), or a physical trait of the user (biometrics). A logical access control system will implement a method to enroll and associate credentials with the user, and then to request that one or more of the user's credentials be authenticated for access to the resource (application, network, device, or operating system). The logical access control system may also log all access attempts for use in auditing who and when someone accessed a specific resource.

Introduction

The key used to open almost any door in the digital realm has traditionally been the password. This was the natural consequence of the fact that somewhere someone manipulated data, from a desktop personal computer (PC) and to prevent this, using passwords began. Furthermore, from a theoretical standpoint, a password can offer extremely strong security since the only place a password needs to be stored is in the user's mind.

In practice however, the mind is a terrible place to store complex secrets; people cannot easily remember complex passwords so they write them down or reveal them to others, and most people end up using the same password everywhere. Exploiting the human factors which affect security is increasingly the quickest path for hackers to break into computer systems. In addition, there are many automated points of attacks on password-based security systems. For instance, a user's password can be compromised via insertion of a hardware or software-based keylogger to trap the keystrokes as they are being entered. And, as computers gain speed, it has become easy to reverse a cryptographic hash, or any other cryptographic representation of a password stored in the computer, even if the password is very complex.

End users do not want to be encumbered with complexities and inconveniences that slow them down while doing their job. On the other hand, businesses increasingly find out that they must implement strong authentication to satisfy industry and government auditors. It is fairly straightforward for a system administrator to patch a piece of software or install a firewall, but it is not trivial to tackle the human factors of security. A secure password policy, such as requiring users to change their passwords every month enforces complexity in construction but in reality makes it more likely that users will find ways to simplify and recall, such as by writing their passwords down on a note under their keyboard. Information technology support costs also go up as more people forget their passwords and need to call the helpdesk. In the end, since passwords are chosen not by the system administrator in a corporation, but by the end users, the system administrator must rely on each user to follow the policy. This typically becomes the weakest link in network security. Other methods, such as tokens and smartcards, succumb to the same challenge – it remains the end user who bears the responsibility of maintaining the security of the credential.

The need to move away from password-based systems can be summarized as follows:

- *Weak passwords are easy to crack.* Most people set their passwords to words or digits they can easily remember, for example, names and birthdays of family members, favorite movie or music stars, and dictionary words. In 2001, a survey of 1,200 British office workers conducted by CentralNic found that almost half chose their own name, a pet's name, or a family member's name as a password. Others based their passwords on celebrity or movie character names, such as "Darth Vader" and "Homer Simpson". Such passwords are easy to crack by guessing or by simple brute force dictionary attacks. Although it is possible, and even advisable, to keep different passwords for different applications and to change them frequently, most people use the same password across different

applications and never change it. Compromising a single password can thus cause a break in security in many applications. For example, a hacker might create a bogus Web site enticing users with freebies if they register with a login name and password. The hacker could then have a good chance of success in using the same login name and password to attack the users' corporate accounts.

- *Strong passwords are difficult to remember.* In an effort to address weak passwords, business often enforce policies to make passwords strong, for example, a business may require that a password is at least 8 characters long, contains at least one digit and one special character, and must be changed every couple of weeks. Such policies backfire. Certainly, longer complex random passwords are more secure, but they are so much harder to remember, which prompts users to write them down in accessible locations such as Post-It notes hidden under the keyboard, an unprotected electronic file on their computer, or other electronic devices such as cellular phones or personal digital assistants (PDAs), creating a security vulnerability. Else, people forget their passwords, which create a financial nightmare to businesses as they have to employ helpdesk support staff to reset forgotten or expired passwords. Cryptographic techniques can provide very long passwords (encryption keys) that the users need not remember; however, these are in turn protected by simple passwords, which defeat their purpose.
- *Password cracking is scalable.* In a password-based network authentication application, a hacker may launch an attack remotely against all the user accounts without knowing any of the users. It costs the hacker almost the same amount of time, effort, and money to attack millions of accounts as it costs to attack one. In fact, the same password (for example, a dictionary word) can be used to launch an attack against (a dictionary of) user accounts. Given that a hacker needs to break only one password among those of all the employees to gain access to a company's intranet, a single weak password compromises the overall security of every system that user has access to. Thus, the entire system's security is only as good as the weakest password.
- *Password and tokens do not provide nonrepudiation.* When a user shares a password with a colleague, there is no way for the system to know who the actual user is. Similarly, tokens can be lost, stolen, shared, duplicated or a hacker could make a master key that opens many locks. Only biometrics can provide a guarantee of authentication that cannot subsequently be refused by a user. It is very hard for the user to deny having accessed a biometric-based system.

Biometrics provide the only credential that does not rely on the end user to maintain its security. Furthermore, biometric systems are potentially cheaper to support and easier to use since the end user does not need to remember complex secrets.

Shrink-wrapped packaged software solutions are available today to enable the use of biometric-based authentication to logon to virtually any consumer and enterprise application, including Microsoft Windows networks, websites, web services, and virtual private networks. Since few applications or operating systems implement native biometric authentication, the role of many such software solutions is to map a successful biometric authentication to the user's long and complex password, which is then used by the application for logon. The end user, however, will likely not need to know his or her underlying password or be able to enter it, and thus, a biometric solution effectively eliminates passwords for the user. Similarly, a user's biometric credential can be bound to the private key associated with a digital certificate to facilitate digital signing of data, such as financial transactions, email, forms, and documents. In addition, to aid compliance the system administrator can access an event log to confirm that a biometric match was performed for access and whether the match was successful or not.

Fingerprint-based solutions, in particular, have emerged as the most common method for logical access control with biometrics. The use of a fingerprint requires the user to declare their credential with a definitive action, such as a finger press or swipe for authentication. Fingerprint readers have attained the size, price, and performance necessary to be integrated in a range of logical access devices, including notebooks, keyboards, mouse, and smartphones.

It is typical for the logical access control applications to have only one user per biometric reader, a reader that may be attached to the user's PC or embedded in her notebook or smartphone. This is unlike most other commercial applications such as physical access control, time and attendance, or authentication at point of sale terminals, where the biometric reader would be shared among many users. Certain logical

access control application deployments may offer the biometric authentication as a choice to the users. A user could chose to use the biometric system or chose to continue using the passwords. In such deployments, the intention of the enterprise is to provide maximum end user convenience while still availing cost savings by reducing helpdesk calls. The above properties of logical access control deployments drive fundamentally different requirements for the single-user biometric reader in terms of accuracy, ease of use, cost, size, and security, as compared to the requirements for the shared-use biometric readers. Shared-use biometric readers traditionally focus on ease of use, durability, and accuracy over a wide demographic population. Single-use biometric readers prioritize low cost, small size, and cryptographic security. For fingerprint-based readers, this trend has manifested itself through the use of placement-based readers for shared-use applications, and swipe-based readers for single-use applications.

Most platforms and peripherals that come with embedded fingerprint readers include software to access the local PC and applications. These applications may include biometric-based access to the PC, pre-boot authentication, full disk encryption, Windows logon, and a general password manager application to facilitate the use of biometric for other applications and websites. Such a suite of applications protects the specific PC on which it is deployed and makes personal access to data more secure, convenient, and fun. Companies such as Dell, Lenovo, Microsoft, and Hewlett-Packard ship platforms and peripherals pre-loaded with such capability. However, these are end user utilities with the scope of use only on the local PC. As a result, they may be challenging and costly to manage if deployed widely in an enterprise since each user will need to setup, enroll his or her biometric, and configure the appropriate policy, all by themselves. Usually the user is given the option to use the biometric system as a cool individual convenience, rather than enforced by an enterprise-wide authentication policy.

The other major class of logical access control biometric application for the enterprise network is server-based solutions. These solutions typically limit the flexibility given to the end user and instead focus on the needs of the organization and the system administrator to deploy, enroll users' biometric credentials into the enterprise directory, and centrally configure enterprise-wide policies. An enterprise-wide policy, however, drives stronger requirements for the reliability, security, and interoperability of the biometric authentication. If it is a business policy that everyone in the organization must use the biometric system for authentication, the reliability of the biometric system must be higher than a ▶ client-side-only solution where the user can opt in to use the biometric system just for convenience. A ▶ server-based logical access control solution generally needs to be interoperable with data coming from many different biometric readers since not every platform in the organization will use the same model of the biometric reader. Interoperability can be accomplished at either the enrollment template level or the biometric image level. Lastly, since a server-based solution typically stores biometric credentials in a central database, the security model of the whole chain from the reader to the server must be considered to protect against hackers and maintain user privacy. However, unlike government deployments that store the user's actual biometric image(s) for archival purposes, a biometric solution used for enterprise authentication typically stores only the biometric enrollment templates.

Biometric systems remove the responsibility of managing credentials from the hands of the end users and therefore resolve the human factors affecting the system security. However, the flip side is that the biometric capture and match process must be trustworthy. Logical access control for users is typically accomplished through a client device, such as a notebook or desktop PC, by authenticating the user to a trusted, managed server. The root challenge of protecting the biometric match process is to remove all means by which a hacker could affect the user authentication by tampering with the client operating system. This can be accomplished by carefully monitoring the health of the client operating system with adequate virus and spyware software, and in the future, with the use of trusted computing, or, if operating from an untrusted client, by removing the client operating system entirely from the system security equation. The practical means to accomplish this is by either performing the biometric match in a secure coprocessor, or by encrypting or digitally signing the raw biometric data on the biometric reader itself so that the biometric data is trusted by the server. Of course, depending on the threats present in a given environment, some deployments of logical access control may need to resolve more than just the human factors of security and will need to use multiple factors of authentication, such as two-factor (biometric and

password) or even three-factor (biometric, smartcard, and PIN) to protect against active adversaries.

After many years of fits and starts as a niche technology, the use of biometrics for logical access control has gained a foothold in protecting corporate assets and networks as the cost of solutions has gone down, and the security and reliability has gone up. Use of biometric authentication for logical access control resolves threats that other secret-based methods such as passwords and tokens cannot, the main threat being the human factors that lower security and are costly and difficult to manage. No security method is a magic bullet, but biometric solutions for logical access control can be a reliable tool or layer to add to a holistic approach to enterprise security.

Specifically, biometrics-based logical access control has found a home in the healthcare and financial industries to help satisfy government compliance directives.

Healthcare

Compliance with the security requirements of the Health Insurance Portability and Accountability Act (HIPAA) of 1996 accelerated the adoption of biometric systems in the U.S. healthcare industry. This regulation does not specify the use of biometrics explicitly, but it states that access to any healthcare data must be restricted through strong user authentication. Such a requirement made the access to healthcare information technology systems and patient data more burdensome. The healthcare industry turned to the biometric systems to get a good balance of convenience, security, and compliance. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) auditing requirements also contributed to the adoption rate.

Once the healthcare industry was educated on the biometric technologies, it adopted biometric systems for other applications as well. Today the healthcare industry uses biometric systems in many different applications to reduce fraud prevalent in the industry and to provide convenience to medical professional without compromising their need for quick and easy access to critical health data. The majority of initial adoption in the healthcare industry was in the employee facing applications. Customer-facing applications have started getting some traction recently. Some examples of business objectives in the healthcare industry that are successfully met with biometrics deployments are:

- Restrict logical access to medical information systems
- Improve hospital efficiency and compliance
- Improve pharmacy efficiency and compliance
- Reduce medical benefits fraud
- Patient verification

Financial

In the U.S., Financial Services Modernization Act of 1999, also known as Gramm-Leach-Bliley Act of 1999 mandates high standards of safeguarding financial transactions, data, and assets. The U.S. Sarbanes-Oxley (SOX) Act of 2002 requires higher security standards for data that is financial or confidential. According to this act, any public company may be liable if it has not taken adequate steps to protect financial records and data. The government considers financial records to be confidential and private. It is imperative that they are secure and access is allowed only to authorized users. Many existing passwords and security policies would not be considered sufficient under SOX. Compliance with these two acts is contributing to an increase in the rate of adoption of biometrics in the financial sector applications. In this respect, the financial industry is somewhat similar to the healthcare industry – adoption of biometric systems in both these industries is being accelerated by government regulations.

Related Entries

▶ Asset Protection
▶ Biometrics Applications, Overview

Access Control, Physical

Colin Soutar
Broadway Avenue, Toronto, ON, Canada

Synonyms

Biometric Readers; Biometric PAC; Physical Access Control

Definition

The use of biometric technologies within physical access control systems is one of the most broadly commercialized sectors of biometrics, outside of forensic applications. A key issue for the successful integration of biometrics within a physical access control system is the interface between the biometric and the access control infrastructures. For this reason, the biometric system must be designed to interface appropriately with a wide range of access control systems. Also, the usability demands of a physical access control system are significant as, typically, all users need to be enrolled for subsequent successful usage more or less on a daily basis. The most significantly-deployed biometric types for access control are: fingerprint; hand geometry; face and iris.

Introduction

The use of biometrics within physical access control (PAC) systems is one of the most broadly commercialized sectors of biometrics, outside of forensic applications. The requirements for the use of biometrics within a larger physical access control system are dependent on the interaction with existing access control infrastructures. For this reason, the biometric system must be designed to interface appropriately with a wide range of access control systems. Also, the usability demands of a physical access control system are significant as all users need to be enrolled for successful usage more or less on a daily basis. The most significantly deployed biometric types for access control are: fingerprint; hand geometry; face and iris. A more recent set of requirements for biometric systems for PAC is that it is also interoperable with logical access control systems – the most broadly recognized example of this requirement is defined in FIPS 201 [1] for access control to federal facilities and computers.

Verification Versus Authorization

As discussed in the introduction, biometric PAC is one of the most commercially deployed applications of biometrics. One of the keys to the success of this application is the capability to interface with multiple PAC systems AND to isolate the act of user verification from the more general PAC system operation of authorization. Achieving these two factors allows a biometric device to be seamlessly added to existing access control systems.

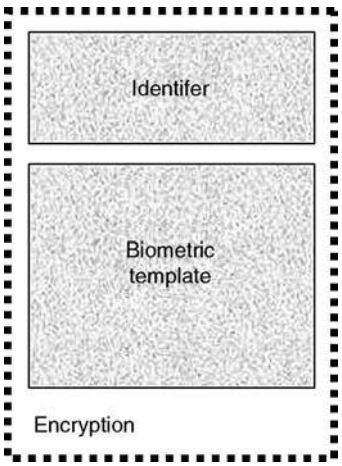
The role that biometric systems serve within the context of a physical access control system is generally to provide evidence (herein referred to as "verification") that an individual is who he/she claims to be. This claim is based on an established persona or *user* that the individual has within the PAC system. It is important to distinguish between the individual's *identity*; an *identifier* (see [2]) by which they are known to a security system – in this case, the PAC system; and the *verification* process which simply verifies that they are the valid owner of the identifier. It is also important to distinguish between authentication (accomplished here via biometric verification) and authorization. Authentication verifies the individual's identity, and authorization permits them to continue with access to the building or facility, based on their status within the PAC system.

As background, consider the various steps comprising the registration of a new user within a PAC system.

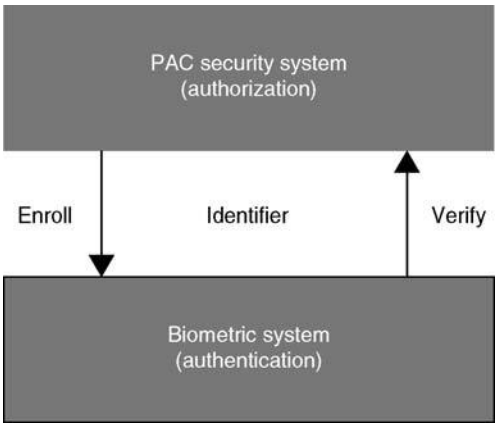
- An administrator of the PAC system will establish the unique identity of the individual. This is typically achieved through the use of so-called "breeder documents" such as employee records, driver's license, passport, etc.
- If the individual is identified as unique, the security system will establish the individual as a new user of the system, and assign a unique identifier by which they are known to the system. An example of an identifier would be the Wiegand data string for physical access control.
- The individual will be instructed to enroll their biometric and the biometric system will create a biometric template that is associated with the user.
- The template will be bound to the identifier, either by physically storing them in related locations in the biometric or security system, or by binding them together using encryption or a digital signature mechanism, to create a user record (see Fig. 1).

Subsequently, when the user requests to access a facility, the following steps are undertaken:

- An individual establishes a claim to the system that he/she is a valid user of the system. This is usually achieved either by inputting the username associated with the user, or by presenting a card



Access Control, Physical. Figure 1 User record, comprising biometric template and user identifier.



Access Control, Physical. Figure 2 Separation of biometric authentication and system authorization.

or other credentials to the system to make the claim.

- The security system ensures that the user record of the claimed user is available to the biometric system (either by transmitting it to the biometric system, or by selecting it within the biometric system), where it will be unbound to produce the template and identifier. Note that as part of the unbinding process either the PAC security system (see Fig. 2) or the biometric system (or both) may verify the authenticity of the user record, by, for example, checking a digital signature.
- The individual is requested to verify that they are the valid owner of the user record, by comparing a live biometric sample with that represented by the template in the user record.
- If a successful match occurs, the identifier that was stored in the user record is relayed to the PAC system where the user is authorized, to rights and privileges according to their PAC security system.

This separation between the authentication of the individual and the authorization of the user is critical for successful integration of biometric systems into general PAC systems. It provides an explicit segregation between the verification process in the biometric system and the rights and privileges that the user is assigned by the security system. This is especially important when considering issues such as the revocation of a user's rights and privileges in a very immediate manner across a wide area system – i.e., a user can still locally verify, but no access action will be permitted as the PAC security system has denied access as a result of the user's authorization privileges having been revoked.

Weigand Format

The most prevalent format for an identifier within a PAC system is the 26-bit Wiegand Format [3]. The 26-bit Wiegand code comprises of 1 parity bit; 8 bits of facility code; 16 bits of identity code; and 1 stop bit. These data thus contains the identifier by which the user is known by a particular access control system. Note that this identifier is explicitly unrelated to the individual's biometric, as described in the previous section. Other formats for identifiers include federal identifiers such as CHUID and FASCN.

Typical Biometrics used for Access Control

Biometrics that are typically used for PAC are those which can provide excellent enrollment rates; throughput rates; and low false rejection rates. The false accept rate is typically set at a rate which is commensurate with the PAC security system requirements, and the false reject rate is thus set by default. Typical biometrics used for PAC are: fingerprint technology; hand geometry; iris technology; and facial recognition. Traditionally, fingerprint and hand geometry have been the

main biometrics used for PAC. As the performance of facial recognition systems improve, for example via dedicated lighting, or by using 3-D surface or texture, this biometric modality is becoming more popular for PAC applications. Similarly, as the cost decreases, and the usability (via verification on the move), of iris recognition systems improves, this modality is also becoming more popular for PAC. Furthermore, systems have been deployed using several of the above biometrics in a combined multi-biometric system.

Interaction with Logical Access Control

As the number of users enrolled in a PAC system that are migrated over to the use of biometrics increases, there is a desire to have the PAC systems interoperable with logical applications systems. This interoperability has several aspects: template interoperability (i.e., it is preferable that the user need not re-enroll for different systems); identifier interoperability (this is especially important where the rights and privileges of the user should span both physical and logical access applications); and event synchronization (for example, a user cannot be granted access to a computer in a room for which they are not authorized to enter). These requirements are more recently being designed into biometric PAC systems; as such PAC systems are required to be a component in a converged physical and logical access control system. A particular example of such as system would be a U.S. Federal system based on HSPD-12, which, in 2004, mandated the establishment of a standard for the identification of Federal employees and contractors, subsequently defined by the Federal Information Processing Standard Publication 201 (FIPS 201), *Personal Identity Verification (PIV) of Federal Employees and Contractors* in February 2005 and *Biometric Data Specification for Personal Identity Verification*, NIST Special Publication 800-76 (SP 800-76). SP 800-76 describe the acquisition and formatting specifications for the biometric credentials of the PIV system and card. In particular, for fingerprints, it calls for compliance to the ANSI/INCITS 378 fingerprint minutiae data interchange format standard for storing two of the captured fingerprints (the left and right index fingers) on the card for use in user verification. This process enables the template interoperability required for a converged physical and logical application.

Access Control, Physical. **Figure 3** Examples of fingerprint and 3-D facial biometric devices for Physical Access Control.

In addition, a unique number stored on the PIV card, known as the CHUID (**C**ard**h**older **U**nique **Id**entifier) is used as the single identifier by which the user is known to both the physical and logical access control systems, thus satisfying the requirement of identifier interoperability as described above.

Related Entries

- ▶ Access Control, Logical
- ▶ Biometric Applications, Overview
- ▶ Biometric System Design
- ▶ Interoperable Performance
- ▶ Multibiometrics

References

1. http://csrc.nist.gov/publications/fips/fips201-1/FIPS-201-1-chng1.pdf
2. Stephen T. Kent., and Lynette I. Millett.: "Ids –Not That Easy." National Academy Press (2002)
3. http://www.hidglobal.com/documents/understandCardDataFormats_wp_en.pdf

Accessibility

- ▶ Ergonomic Design for Biometric System

ACE-V

ACE-V is the four phase identification protocol which includes analysis, comparison, evaluation, and verification.

► Earprints, Forensic Evidence of

Action Categorization

► Psychology of Gait and Action Recognition

Action Understanding

► Psychology of Gait and Action Recognition

Active (Contour, Shape, Appearance) Models

► Deformable Models

Acupuncture

Acupuncture is a treatment where sharp, thin needles are inserted in the body at very specific points. It is one of the forms of treatment in traditional Chinese medicine.

► Skull, Forensic Evidence of

AdaBoost

AdaBoost (short for Adaptive Boosting) is a machine learning algorithm that learns a strong classifier by combining an ensemble of weak (moderately accurate) classifiers with weights. The discrete AdaBoost algorithm was originally developed for classification using the exponential loss function and is an instance within the boosting family. Boosting algorithms can also be derived from the perspective of function approximation with gradient descent and applications for regression.

► Face Detection

Adapted Fusion

Adapted fusion in the framework of multi-biometric score fusion refers to the techniques in which a baseline fusion function is first constructed based on some general knowledge of the problem at hand, and then adjusted during the operation of the system. The adaptation can be based on ancillary information such as: the user being claimed (adapted user-specific fusion), quality measures of the input biometrics (► quality-based fusion), or other kind of environmental information affecting the various information channels being fused.

► Fusion, User-Specific

Adaptive Learning

► Incremental Learning

Affective Computing

The research area concerned with computing that relates to, arises from, or deliberately influences

emotion. Specifically, the research area of machine analysis of human affective states and employment of this information to build more natural, flexible (affective) user interfaces goes by a general name of affective computing. Affective computing expands HCI by including emotional communication together with appropriate means of handling affective information.

► Facial Expression Recognition

Albedo

For a reflecting surface, Albedo is the fraction of the incident light that is reflected. This is a summary characteristic of the surface. Reflection can be quite complicated and a complete description of the reflectance properties of a surface requires the specification of the bidirectional reflectance distribution function as a function of wavelength and polarization for the surface.

► Iris Device

Alignment

Alignment is the process of transforming two or more sets of data into a common coordinate system. For example, two fingerprint scans acquired at different times each belong to their own coordinate system; this is because of rotation, translation, and non-linear distortion of the finger. In order to match features between the images, a correspondence has to be established. Typically, one image (signal) is referred to as the reference and the other image is the target, and the goal is to map the target onto the reference. This transformation can be both linear and nonlinear based on the deformations undergone during acquisition. Position-invariant features, often used to avoid registration, face other concerns like robustness to local variation such as non-linear distortions or occlusion.

► Biometric Algorithms

Altitude

Altitude is the angle between a line that crosses a plane and its projection on it, ranging from 0° (if the line is contained in the plane) to 90° (if the line is orthogonal to the plane). This measure is a component of the pen orientation in handwriting capture devices.

► Signature Features

Ambient Space

The space in which the input data of a mathematical object lie, for example, the plane for lines.

► Manifold Learning

American National Standards Institute (ANSI)

ANSI is a non-government organization that develops and maintains voluntary standards for a wide range of products, processes, and services in the United States. ANSI is a member of the international federation of standards setting bodies, the ISO.

► Iris Device

Anatomy

It is a branch of natural science concerned with the study of the bodily structure of living beings, especially as revealed by dissection. The word "Anatomy"

originates from the Old French word "Anatomie," or a Late Latin word "Anatmia." Anatomy implies, "ana" meaning "up" and "tomia" meaning "cutting."

► Anatomy of Face
► Anatomy of Hand

Anatomy of Eyes

Kristina Irsch, David L. Guyton
The Wilmer Ophthalmological Institute, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

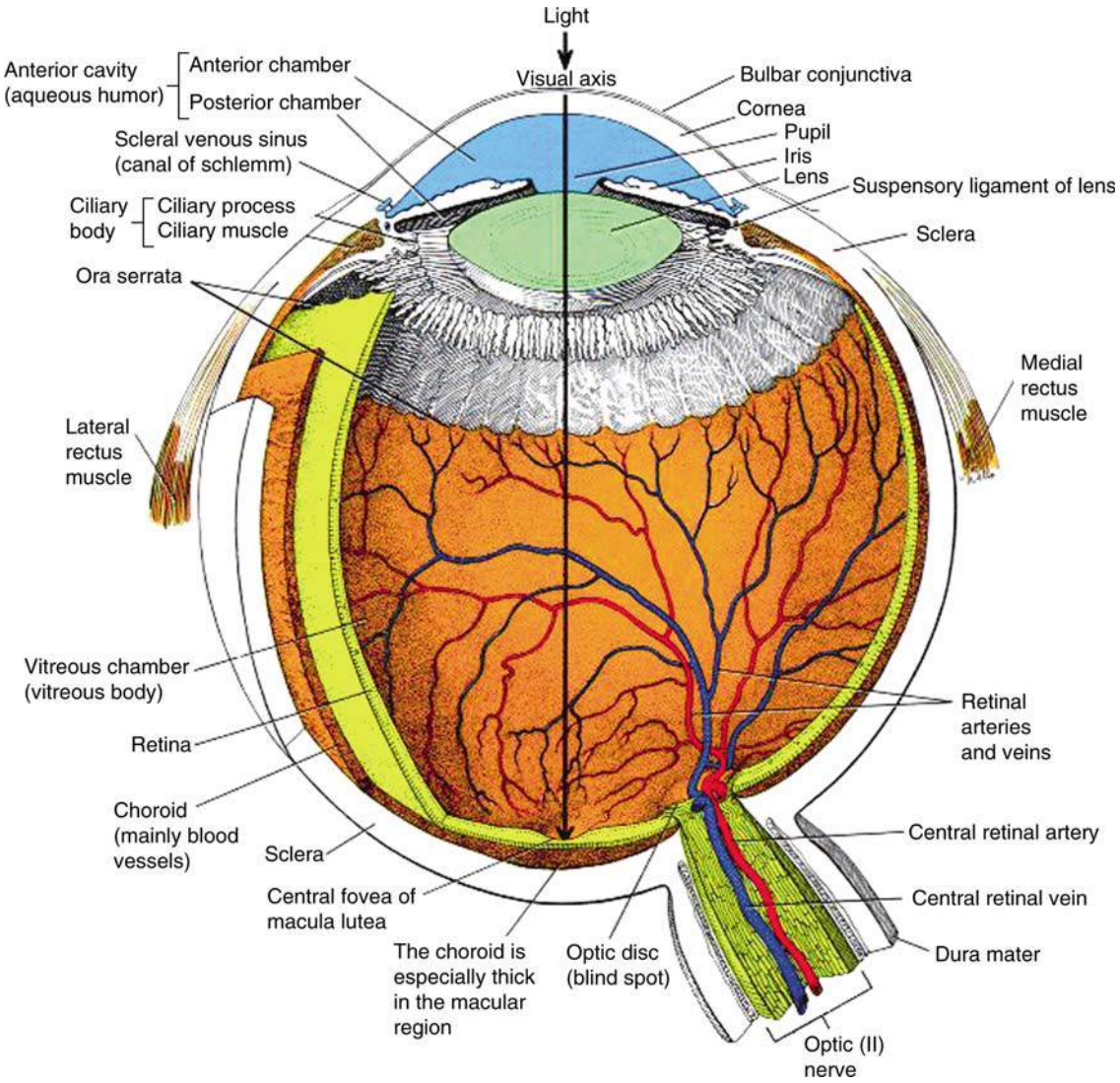
Definition

The human eye is one of the most remarkable sensory systems. Leonardo da Vinci was acutely aware of its prime significance: "The eye, which is termed the window of the soul, is the chief organ whereby the *senso comune* can have the most complete and magnificent view of the infinite works of nature" [1]. Human beings gather most of the information about the external environment through their eyes and thus rely on sight more than on any other sense, with the eye being the most sensitive organ we have. Besides its consideration as a window to the soul, the eye can indeed serve as a window to the identity of an individual. It offers unique features for the application of identification technology. Both the highly detailed texture of the iris and the fundus blood vessel pattern are unique to every person, providing suitable traits for biometric recognition.

Anatomy of the Human Eye

The adult eyeball, often referred to as a spherical globe, is only approximately spherical in shape, with its largest diameter being 24 mm antero-posteriorly [2, 3]. A schematic drawing of the human eye is shown in Fig. 1. The anterior portion of the eye consists of the cornea, iris, pupil, and crystalline lens. The pupil serves as an aperture which is adjusted by the surrounding ► iris, acting as a diaphragm that regulates the amount of light entering the eye. Both the iris and the pupil are covered by the convex transparent cornea, the major refractive component of the eye due to the huge difference in refractive index across the air-cornea interface [5]. Together with the crystalline lens, the cornea is responsible for the formation of the optical image on the retina. The crystalline lens is held in place by suspensory ligaments, or zonules, that are attached to the ciliary muscle. Ciliary muscle actions cause the zonular fibers to relax or tighten and thus provide accommodation, the active function of the crystalline lens. This ability to change its curvature, allowing objects at various distances to be brought into sharp focus on the retinal surface, decreases with age, with the eye becoming "presbyopic." Besides the cornea and crystalline lens, both the vitreous and aqueous humor contribute to the dioptric apparatus of the eye, leading to an overall refractive power of about 60 diopters [3]. The aqueous humor fills the anterior chamber between the cornea and iris, and also fills the posterior chamber that is situated between the iris and the zonular fibers and crystalline lens. Together with the vitreous humor, or vitreous, a loose gel filling the cavity between the crystalline lens and retina, the aqueous humor is responsible for maintaining the intraocular pressure and thereby helps the eyeball maintain its shape. Moreover, this clear watery fluid nourishes the cornea and crystalline lens. Taken all together, with its refracting constituents, self-adjusting aperture, and finally, its detecting segment, the eye is very similar to a photographic camera. The film of this optical system is the ► retina, the multilayered sensory tissue of the posterior eyeball onto which the light entering the eye is focused, forming a reversed and inverted image. External to the retina is the *choroid*, the layer that lies between retina and sclera. The choroid is primarily composed of a dense capillary plexus, as well as small arteries and veins [5]. As it consists of numerous blood vessels and thus contains many blood cells, the choroid supplies most of the back of the eye with necessary oxygen and nutrients. The sclera is the external fibrous covering of the eye. The visible portion of the sclera is commonly known as the "white" of the eye.

Both iris and retina are described in more detail in the following sections due to their major role in biometric applications.



Anatomy of Eyes. **Figure 1** Schematic drawing of the human eye [4].

Iris

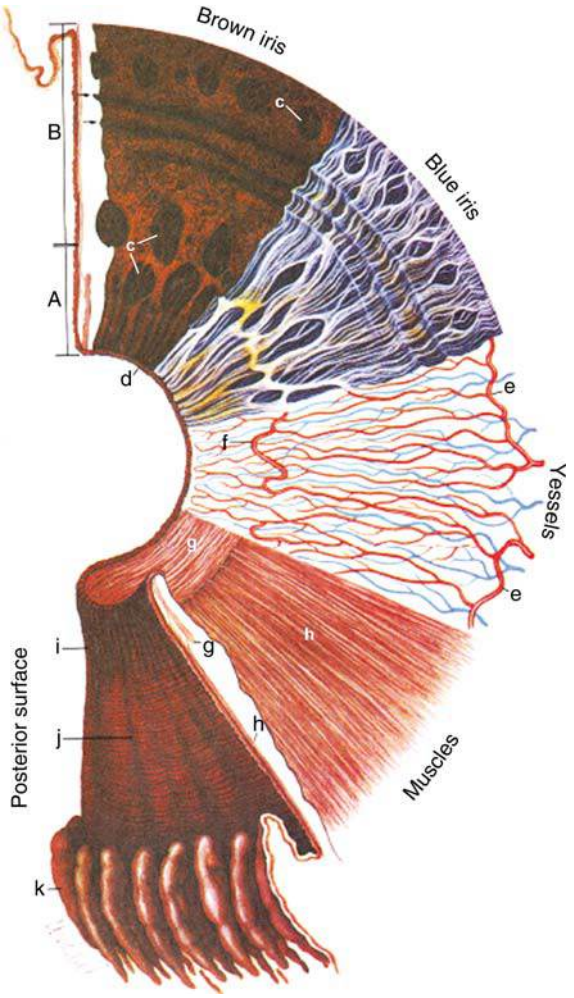
The iris may be considered as being composed of four different layers [3], starting from anterior to posterior: (1) *Anterior border layer* which mainly consists of fibroblasts and pigmented melanocytes, interrupted by large, pit-like holes, the so-called crypts of Fuchs; (2) S*troma* containing loosely arranged collagen fibers that are condensed around blood vessels and nerve fibers. Besides fibroblasts and melanocytes, as present in the previous layer, clump cells and mast cells are found in the iris stroma. It is the pigment in the melanocytes that determines the color of the iris, with blue eyes representing a lack of melanin pigment. The *sphincter pupillae muscle*, whose muscle fibers encircle the pupillary margin, lies deep inside the stromal layer. By contracting, the sphincter causes pupil constriction, which subsequently results in so-called contraction furrows in the iris. These furrows deepen with dilation of the pupil, caused by action of the dilator muscle, which is formed by the cellular processes of the (3) *Anterior epithelium*. The dilator pupillae muscle belongs to the anterior epithelial layer, with its cells being myoepithelial [6]. Unlike the sphincter muscle,

the muscle fibers of the dilator muscle are arranged in a radial pattern, terminating at the iris root; (4) *Posterior pigmented epithelium* whose cells are columnar and more heavily pigmented in comparison with the anterior epithelial cells. The posterior epithelial layer functions as the main light absorber within the iris.

A composite view of the iris surfaces and layers is shown in Fig. 2, which indicates the externally visible iris features, enhancing the difference in appearance between light and dark irides (iris features and anatomy). Light irides show more striking features in visible light because of higher contrast. But melanin is relatively transparent to near-infrared light, so viewing the iris with light in the near-infrared range will uncover deeper features arising from the posterior layers, and thereby reveals even the texture of dark irides that is often hidden with visible light.

In general, the iris surface is divided into an inner pupillary zone and an outer ciliary zone. The border between these areas is marked by a sinuous structure, the so-called collarette. In addition to the particular arrangement of the iris crypts themselves, the structural features of the iris fall into two categories [7]: (1) Features that relate to the pigmentation of the iris (e.g., pigment spots, pigment frill), and (2) movement-related features, in other words features of the iris relating to its function as pupil size control (e.g., iris sphincter, contraction furrows, radial furrows).

Among the visible features that relate to the pigmentation belong small elevated white or yellowish Wölfflin spots in the peripheral iris, which are predominantly seen in light irides [3]. The front of the iris may also reveal iris freckles, representing random accumulations of melanocytes in the anterior border layer. Pigment frill or pupillary ruff is a dark pigmented ring at the pupil margin, resulting from a forward extension of the posterior epithelial layer. In addition to the crypts of Fuchs, predominantly occurring adjacent to the collarette, smaller crypts are located in the periphery of the iris. These depressions, that are dark in appearance because of the darkly pigmented posterior layers, are best seen in blue irides. Similarly, a buff-colored, flat, circular strap-like muscle becomes apparent in light eyes, that is, the iris sphincter. The contraction furrows produced when it contracts, however, are best noticeable in dark irides as the base of those concentric lines is less pigmented. They appear near the outer part of the ciliary zone, and are



Anatomy of Eyes. **Figure 2** Composite view of the surfaces and layers of the iris. Crypts of Fuchs (**c**) are seen adjacent to the collarette in both the pupillary (**a**) and ciliary zone (**b**). Several smaller crypts occur at the iris periphery. Two arrows (top left) indicate circular contraction furrows occurring in the ciliary area. The pupillary ruff (**d**) appears at the margin of the pupil, adjacent to which the circular arrangement of the sphincter muscle (**g**) is shown. The muscle fibers of the dilator (**h**) are arranged in a radial fashion. The last sector at the bottom shows the posterior surface with its radial folds (**i** and **j**). (Reproduced with permission from [5]).

crossed by radial furrows occurring in the same region. Posterior surface features of the iris comprise structural and circular furrows, pits, and contraction folds. The latter, for instance, also known as Schwalbe's contraction

folds, cause the notched appearance of the pupillary margin.

All the features described above contribute to a highly detailed iris pattern that varies from one person to the next. Even in the same individual, right and left irides are different in texture. Besides its uniqueness, the iris is a protected but readily visible internal organ, and it is essentially stable over time [7, 8]. Thus the iris pattern provides a suitable physical trait to distinguish one person from another. The idea of using the iris for biometric identification was originally proposed by the ophthalmologist Burch in 1936 [9]. However, it took several decades until two other ophthalmologists, Flom and Safir [7], patented the general concept of iris-based recognition. In 1989, Daugman, a mathematician, developed efficient algorithms for their system [8–10]. His mathematical formulation provides the basis for most iris scanners now in use. Current iris recognition systems use infrared-sensitive video cameras to acquire a digitized image of the human eye with near-infrared illumination in the 700−900 nm range. Then image analysis algorithms extract and encode the iris features into a binary code which is stored as a template. Elastic deformations associated with pupil size changes are compensated for mathematically. As pupil motion is limited to living irides, small distortions are even favorable by providing a control against fraudulent artificial irides [8, 10].

Imaging the iris with near-infrared light not only greatly improves identification in individuals with very dark, highly pigmented irides, but also makes the system relatively immune to anomalous features related to changes in pigmentation. For instance, melanomas/tumors may develop on the iris and change its appearance. Furthermore, some eye drops for glaucoma treatment may affect the pigmentation of the iris, leading to coloration changes or pigment spots. However, as melanin is relatively transparent to near-infrared light and basically invisible to monochromatic cameras employed by current techniques of iris recognition, none of these pigment-related effects causes significant interference [9, 10].

Retina

As seen in an ordinary histologic cross-section, the retina is composed of distinct layers. The retinal layers from the vitreous to choroid [2, 3] are: (1) *Internal limiting membrane*, formed by both retinal and vitreal elements [2]; (2) *Nerve fiber layer*, which contains the axons of the ganglion cells. These nerve fibers are bundled together and converge to the optic disc, where they leave the eye as the optic nerve. The cell bodies of the ganglion cells are situated in the (3) *ganglion cell layer*. Numerous dendrites extend into the (4) *inner plexiform layer* where they form synapses with interconnecting cells, whose cell bodies are located in the (5) *inner nuclear layer*; (6) *Outer plexiform layer*, containing synaptic connections of photoreceptor cells; (7) *Outer nuclear layer*, where the cell bodies of the photoreceptors are located; (8) *External limiting membrane*, which is not a membrane in the proper sense, but rather comprises closely packed junctions between photoreceptors and supporting cells. The photoreceptors reside in the (9) *receptor layer*. They comprise two types of receptors: rods and cones. In each human retina, there are 110–125 million rods and 6.3–6.8 million cones [2]. Light contacting the photoreceptors and thereby their light-sensitive photopigments, are absorbed and transformed into electrical impulses that are conducted and further relayed to the brain via the optic nerve; (10) *Retinal pigment epithelium*, whose cells supply the photoreceptors with nutrients. The retinal pigment epithelial cells contain granules of melanin pigment that enhance visual acuity by absorbing the light not captured by the photoreceptor cells, thus reducing glare. The most important task of the retinal pigment epithelium is to store and synthesize vitamin A, which is essential for the production of the visual pigment [3]. The pigment epithelium rests on Bruch's membrane, a basement membrane on the inner surface of the choroid.

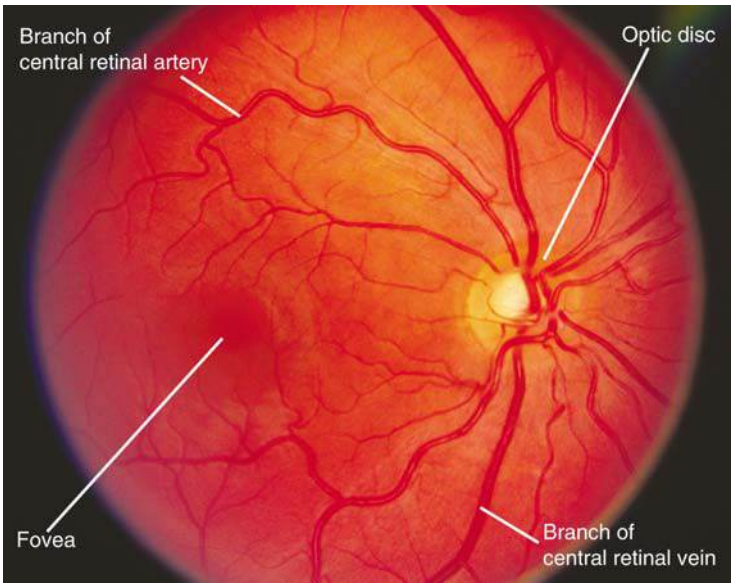
There are two areas of the human retina that are structurally different from the remainder, namely the ▶ fovea and the optic disc. The fovea is a small depression, about 1.5 mm across, at the center of the macula, the central region of the retina [11]. There, the inner layers are shifted aside, allowing light to pass unimpeded to the photoreceptors. Only tightly packed cones, and no rods, are present at the foveola, the center of the fovea. There are also more ganglion cells accumulated around the foveal region than elsewhere. The fovea is the region of maximum visual acuity.

The optic disc is situated about 3 mm (15 degrees of visual angle) to the nasal side of the macula [11]. It contains no photoreceptors at all and hence is responsible for the blind spot in the field of vision. Both choroidal capillaries and the central retinal artery and vein supply

the retina with blood. A typical fundus photo taken with visible light of a healthy right human eye is illustrated in Fig. 3, showing the branches of the central artery and vein as they diverge from the center of the disc. The veins are larger and darker in appearance than the arteries. The temporal branches of the blood vessels arch toward and around the macula, seen as a darker area compared with the remainder of the fundus, whereas the nasal vessels course radially from the nerve head. Typically, the central ▶ retinal blood vessels divide into two superior and inferior branches, yielding four arterial and four venous branches that emerge from the optic disc. However, this pattern varies considerably [6]. So does the choroidal blood vessel pattern, forming a matting behind the retina, which becomes visible when observed with light in the near-infrared range [12]. The blood vessels of the choroid are even apparent in the foveal area, whereas retinal vessels rarely occur in this region.

In the 1930s, Simon and Goldstein noted that the blood vessel pattern is unique to every eye. They suggested using a photograph of the retinal blood vessel pattern as a new scientific method of identification [13]. The uniqueness of the pattern mainly comprises the number of major vessels and their branching characteristics. The size of the optic disc also varies across individuals. Because this unique pattern remains essentially unchanged throughout life, it can potentially be used for biometric identification [12, 14].

Commercially available retina scans recognize the blood vessels via their light absorption properties. The original Retina Scan used green light to scan the retina in a circular pattern centered on the optic nerve head [14]. Green light is strongly absorbed by the dark red blood vessels and is somewhat reflected by the retinal tissue, yielding high contrast between vessels and tissue. The amount of light reflected back from the retina was detected, leading to a pattern of discontinuities, with each discontinuity representing an absorbed spot caused by an encountered blood vessel during the circular scan. To overcome disadvantages caused by visible light, such as discomfort to the subject and pupillary constriction decreasing the signal intensity, subsequent devices employ near-infrared light instead. The generation of a consistent signal pattern for the same individual requires exactly the same alignment/fixation of the individual's eye every time the system is used. To avoid variability with head tilt, later designs direct the scanning beam about the visual axis, therefore centered on the fovea, so that the captured vascular patterns are more immune to head tilt [12]. As mentioned before, the choroidal vasculature forms a matting behind the retina even in the region of the



Anatomy of Eyes. **Figure 3** Fundus picture of a right human eye.

macula and becomes detectable when illuminated with near-infrared light. Nevertheless the requirement for steady and accurate fixation still remains a problem because if the eye is not aligned exactly the same way each time it is measured, the identification pattern will vary. Reportedly a more recent procedure solves the alignment issue [15]. Instead of using circular scanning optics as in the prior art, the fundus is photographed, the optic disc is located automatically in the obtained retinal image, and an area of retina is analyzed in fixed relationship to the optic disc.

Related Entries

- ▶ Iris Acquisition Device
- ▶ Iris Device
- ▶ Iris Recognition, Overview
- ▶ Retina Recognition

References

1. Pevsner, J.: Leonardo da Vinci's contributions to neuroscience. Trends Neurosci. **25**, 217–220 (2002)
2. Davson, H.: The Eye, vol. 1a, 3rd edn. pp. 1–64. Academic Press, Orlando (1984)
3. Born, A.J., Tripathi, R.C., Tripathi, B.J.: Wolff's Anatomy of the Eye and Orbit, 8th edn. pp. 211–232, 308−334, 454−596. Chapman & Hall Medical, London (1997)
4. Ian Hickson's Description of the Eye. http://academia.hixie.ch/bath/eye/home.html. Accessed (1998)
5. Warwick, R., Williams, P.L. (eds.): Gray's Anatomy, 35th British edn. pp. 1100–1122. W.B Saunders, Philadelphia (1973)
6. Oyster, C.W.: The Human Eye: Structure and Function, pp. 411–445, 708−732. Sinauer Associates, Inc., Sunderland (1999)
7. Flom, L., Safir, A.: Iris recognition system, US Patent No. 4,641,349 Feb (1987)
8. Daugman, J.: Biometric personal identification system based on iris analysis, US Patent No. 5,291,560 Mar (1994)
9. Daugman, J.: Iris Recognition. Am. Sci. **89**, 326–333 (2001)
10. Daugman, J.: Recognizing persons by their iris patterns. In: Biometrics: Personal Identification in Networked Society. Online textbook, http://www.cse.msu.edu/~cse891/Sect601/textbook/5.pdf. Accessed August (1998)
11. Snell, R.S., Lemp, M.A.: Clinical Anatomy of the Eye, pp. 169–175. Blackwell, Inc., Boston (1989)
12. Hill, R.B.: Fovea-centered eye fundus scanner, US Patent No. 4,620,318 Oct (1986)
13. Simon, C., Goldstein, I.: A new scientific method of identification. NY State J Med **35**, 901–906 (1935)
14. Hill, R.B.: Apparatus and method for identifying individuals through their retinal vasculature patterns, US Patent No. 4,109,237 Aug (1978)
15. Marshall, J., Usher, D.: Method for generating a unique and consistent signal pattern for identification of an individual, US Patent No. 6,757,409 Jun (2004)

Anatomy of Face

Anne M. Burrows[1], Jeffrey F. Cohn[2]
[1]Duquesne University, Pittsburgh, PA, USA
[2]University of Pittsburgh, Pittsburgh, PA, USA

Synonyms

Anatomic; Structural and functional anatomy

Definition

Facial anatomy – The soft-tissue structures attached to the bones of the facial skeleton, including epidermis, dermis, subcutaneous fascia, and mimetic musculature.

Introduction

Face recognition is a leading approach to person recognition. In well controlled settings, accuracy is comparable to that of historically reliable biometrics including fingerprint and iris recognition [1]. In less-controlled settings, accuracy is attenuated with variation in pose, illumination, and facial expression among other factors. A principal research challenge is to increase robustness to these sources of variation, and to improve performance in unstructured settings in which image acquisition may occur without active subject involvement.

Current approaches to face recognition are primarily data driven. Use of domain knowledge tends to be limited to the search for relatively stable facial features, such as the inner canthi and the philtrum for image alignment, or the lips, eyes, brows, and face contour for feature extraction. More explicit reference to domain knowledge

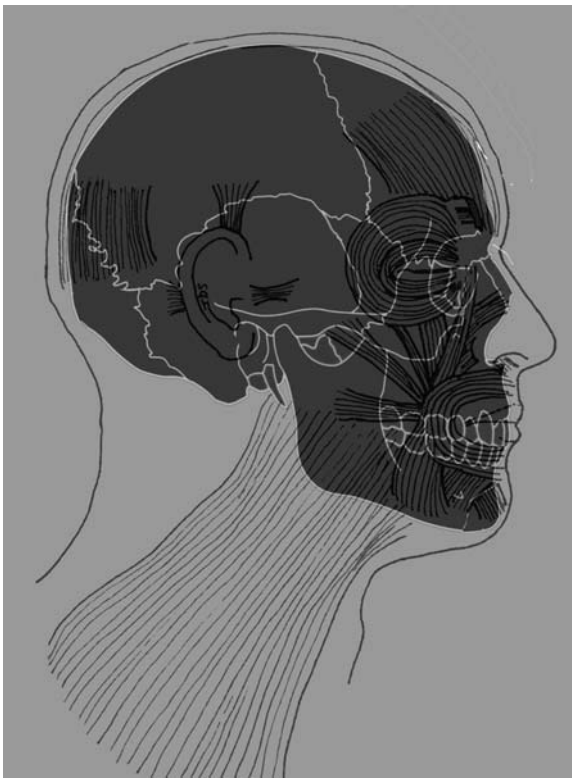
of the face is relatively rare. Greater use of domain knowledge from facial anatomy can be useful in improving the accuracy, speed, and robustness of face recognition algorithms. Data requirements can be reduced, since certain aspects need not be inferred, and parameters may be better informed. This chapter provides an introduction to facial ► anatomy that may prove useful towards this goal. It emphasizes facial skeleton and musculature, which bare primary responsibility for the wide range of possible variation in face identity.

Morphological Basis for Facial Variation Among Individuals

The Skull

It has been suggested that there is more variation among human faces than in any other mammalian species except for domestic dogs [2]. To understand the factors responsible for this variation, it is first necessary to understand the framework of the face, the skull. The bones of the skull can be grouped into three general structural regions: the *dermatocranium*, which surrounds and protects the brain; the *basicranium*, which serves as a stable platform for the brain; and the *viscerocranium* (facial skeleton) which houses most of the special sensory organs, the dentition, and the oronasal cavity [3]. The facial skeleton also serves as the bony framework for the ► mimetic musculature. These muscles are stretched across the facial skeleton like a mask (Fig. 1). They attach into the dermis, into one another, and onto facial bones and nasal cartilages. Variation in facial appearance and expression is due in great part to variation in the facial bones and the skull as a whole [2].

The viscerocranium (Fig. 2) is composed of 6 paired bones: the maxilla, nasal, zygomatic (malar), lacrimal, palatine, and inferior nasal concha. The vomer is a midline, unpaired bone; and the mandible, another unpaired bone, make up the 13th and 14th facial bones [3]. While not all of these bones are visible on the external surface of the skull, they all participate in producing the ultimate form of the facial skeleton. In the fetal human there are also paired premaxilla bones, which fuse with the maxilla sometime during the late fetal or early infancy period [2]. Separating the bones from one another are sutures. Facial sutures

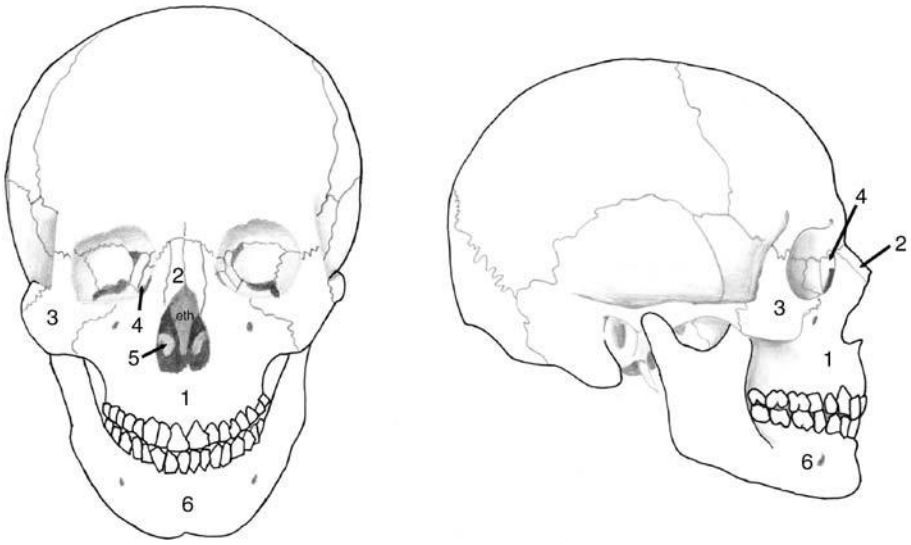


Anatomy of Face. Figure 1 Mimetic musculature and underlying facial skeleton. © Tim Smith.

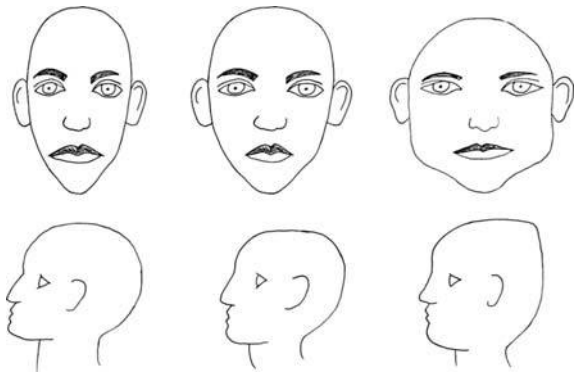
are fairly immobile fibrous joints that participate in the growth of the facial bones, and they absorb some of the forces associated with chewing [2]. Variation in the form of these bones is the major reason that people look so different [4].

While there are many different facial appearances, most people fall into one of three types of head morphologies: ► *dolicocephalic*, meaning a long, narrow head with a protruding nose (producing a leptoproscopic face); ► *mesocephalic*, meaning a proportional length to width head (producing a mesoproscopic face); and *brachycephalic*, meaning a short, wide head with a relatively abbreviated nose (producing a euryproscopic face) (Fig. 3).

What accounts for this variation in face shape? While numerous variables are factors for this variation, it is largely the form of the cranial base that establishes overall facial shape. The facial skeleton is attached to the cranial base which itself serves as a template for establishing many of the angular, size-related, and



Anatomy of Face. **Figure 2** Frontal view (left) and side view (right) of a human skull showing the bones that make up the facial skeleton, the viscerocranium. Note that only the bones that compose the face are labeled here. Key: 1 – maxilla, 2 – nasal, 3 – zygomatic (malar), 4 – lacrimal, 5 – inferior nasal concha, 6 – mandible. The vomer is not shown here as it is located deeply within the nasal cavity, just inferior to the ethmoid (eth). While the maxilla is shown here as a single bone it remains paired and bilateral through the 20's and into the 30's [2]. The mandible is shown here as an unpaired bone as well. It begins as two separate dentaries but fuses into a single bone by 6 months of age [2]. Compare modern humans, *Homo sapiens*, with the fossil humans in Fig. 6, noting the dramatic enlargement of the brain and reduction in the "snout". © Anne M. Burrows.



Anatomy of Face. **Figure 3** Representative human head shapes (top row) and facial types (bottom row). Top left – dolicocephalic head (long and narrow); middle – mesocephalic head; right – brachycephalic head (short and wide). Bottom left – leptoproscopic face (sloping forehead, long, protuberant nose); middle – mesoproscopic face; right – euryproscopic face (blunt forehead with short, rounded nose). © Anne M. Burrows

topographic features of the face. Thus a dolicocephalic cranial base sets up a template for a long, narrow face while a brachycephalic cranial base sets up a short, wide face. A soft-tissue facial mask stretched over each of these facial skeleton types must reflect the features of the bony skull. While most human population fall into a brachycephalic, mesocephalic, or dolicocephalic head shape, the variation in shape within any given group typically exceeds variation between groups [2]. Overall, though, dolicocephalic forms tend to predominate in the northern and southern edges of Europe, the British Isles, Scandinavia, and sub-Saharan Africa. Brachycephalic forms tend to predominate in central Europe and China and mesocephalic forms tend to be found in Middle Eastern countries and various parts of Europe [4]. Geographic variation relates to relative genetic isolation of human population following dispersion from Africa approximately 50,000 years ago.

Variation in facial form is also influenced by sex, with males tending to have overall larger faces. This

dimorphism is most notable in the nose and forehead. Males, being larger, need more air in order to support larger muscles and viscera. Thus, the nose as the entrance to the airway will be longer, wider, and more protrusive with flaring nostrils. This larger nose is associated with a more protrusive, sloping forehead while female foreheads tend to be more upright and bulbous. If a straight line is drawn in profile that passes vertically along the surface of the upper lip, the female forehead typically lies far behind the line with only the tip of the nose passing the line. Males, on the other hand, tend to have a forehead that is closer to the line and have more of the nose protruding beyond the line [2, 5]. The protruding male forehead makes the eyes appear to be deeply set with less prominent cheek bones than in females. Because of the less protrusive nose and forehead the female face appears to be flatter than that of male's. Males are typically described as having deep and topographically irregular faces.

What about the variation in facial form with change in age? Facial form in infants tends to be brachycephalic because the brain is precocious relative to the face, which causes the dermatocranium and basicranium to be well-developed relative to the viscerocranium. As people age to adulthood, the primary cue to the aging face is the sagging soft-tissue: the ► collagenous fibers and ► proteoglycans of the dermis decline in number such that dehydration occurs. Additionally, subcutaneous fat deposits tend to be reabsorbed, which combined with dermal changes yields a decrease in facial volume, skin surplus (sagging of the skin), and wrinkling [4].

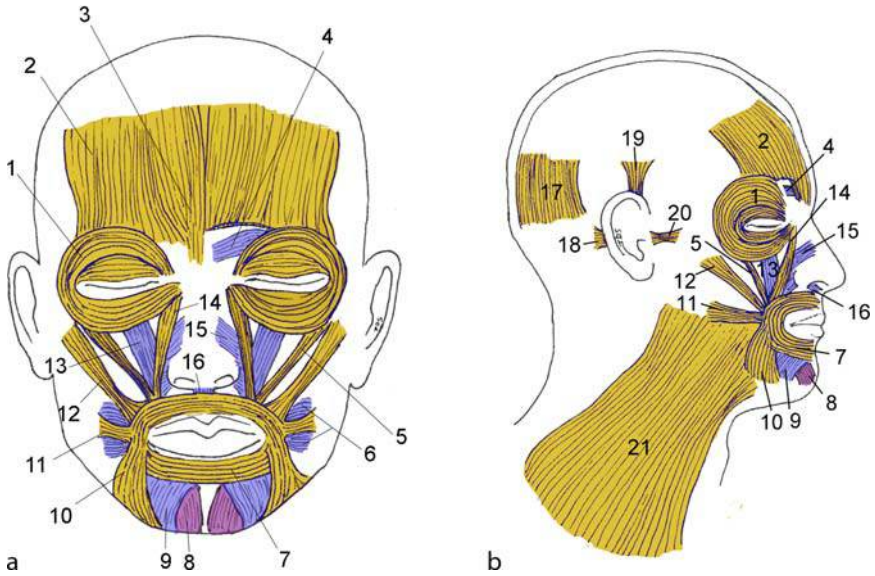
Musculature and Associated Soft Tissue

Variation in facial appearance among individuals is also influenced by the soft tissue structures of the facial skeleton: the mimetic musculature, the superficial ► fasciae, and adipose deposits. All humans generally have the same mimetic musculature (Fig. 4). However, this plan does vary. For instance, the risorious muscle, which causes the lips to flatten and stretch laterally, was found missing in 22 of 50 specimens examined [6]. Recent work [7, 8] has shown that the most common variations involve muscles that are nonessential for making five of the six universal facial expressions of emotion (fear, anger, sadness, surprise, and happiness). The sixth universal facial expression, disgust can be formed from a variety of different muscle combinations, so there are no 'essential' muscles. The most variable muscles are the risorius, depressor septi, zygomaticus minor, and procerus muscles. Muscles that vary the least among individuals were found to be the orbicularis oris, orbicularis occuli, zygomaticus major, and depressor anguli oris muscles, all of which are necessary for creating the aforementioned universal expressions.

In addition to presence, muscles may vary in form, location, and control. The bifid, or double, version of the zygomaticus major muscle has two insertion points rather than the more usual single insertion point. The bifid version causes dimpling or a slight depression to appear when the muscle contracts [6, 9, 10]. The platysma muscle inserts in the lateral cheek or on the skin above the inferior margin of the mandible. Depending on insertion region, lateral furrows are formed in the cheek region when the muscle contracts. Muscles also vary in the relative proportion of slow to fast twitch fibers. Most of this variation is between muscles. The orbicularis occuli and zygomaticus major muscles, for instance, have relatively high proportions of fast twitch fibers relative to some other facial muscles [11]. For the orbicularis oculi, fast twitch fibers are at least in part an adaptation for eye protection. Variation among individuals in the ratio of fast to slow twitch fibers is relatively little studied, but may be an important source of individual difference in facial dynamics. Overall, the apparent predominance of fast-twitch fibers in mimetic musculature indicates a muscle that is primarily capable of producing a quick contraction but one that fatigues quickly (slow-twitch fibers give a muscle a slow contraction speed but will not fatigue quickly). This type of contraction is consistent with the relatively fast neural processing time for facial expression in humans [8].

A final source of variation is cultural. Facial movements vary cross-culturally [12] but there is little literature detailing racial differences in mimetic muscles. To summarize, variation in presence, location, form, and control of facial muscles influences the kind of facial movement that individuals create. Knowledge of such differences in expression may be especially important when sampling faces in the natural environment in which facial expression is common.

While there are no studies detailing individual variation in the other soft tissue structures of the face,



Anatomy of Face. **Figure 4** Human mimetic musculature in (A.) frontal and (B.) right side views. Key: 1 – orbicularis occuli m., 2 – frontalis m., 3 – procerus m., 4 – corrugator supercilli m., 5 – zygomaticus minor m., 6 – buccinator m., 7 – orbicularis oris m., 8 – mentalis m., 9 – depressor labii inferioris m., 10 – depressor anguli oris m., 11 – risorius m., 12 – zygomaticus major m., 13 – levator labii superioris m., 14 – levator labii superioris alaeque nasi m., 15 – nasalis m., 16 – depressor septi m., 17 – occipitalis m., 18 – posterior auricularis m., 19 – superior auricularis m., 20 – anterior auricularis m., 21 – platysma m. Color coding represents depth of musculature with muscles colored yellow being the most superficial, muscles colored blue being intermediate in depth, and muscles colored purple being the deepest. Note that the buccinator m. (#6) is not considered to be a mimetic muscle but it is included here as a muscle located on the face that is innervated by the facial nerve [7]. © Anne M. Burrows.

they may also affect facial appearance. The facial soft-tissue architecture is a layered arrangement with the *epidermis* and *dermis* being most superficial, followed by the subcutaneous fat, superficial *fascia*, mimetic musculature, and deep facial fascia (such as the parotid/masseteric fascia) and the buccal fat pad [13]. The superficial fascia mainly consists of the *SMAS* (the superficial musculoaponeurotic system). This is a continuous fibromuscular fascia found in the face that invests and interlocks the mimetic muscles. It sweeps over the parotid gland, up to the zygomatic arch, across the cheeks and lips and down to the region of the platysma muscle. This sheet is also attached to the deep fascia of the face and the dermis [13]. The collagen fibers found throughout the SMAS deteriorate with age, contributing to the sagging facial appearance during the aging process. In addition, fat deposits in the facial region, especially the buccal fat pad located between the masseter muscle and the orbicularis oris muscle, also break down with age and contributes to the sagging [13].

Contributing to change with age are the cumulative effects of individual differences in facial expression. When facial muscles contract, facial lines and furrows appear parallel to the direction of the contraction. With aging, the elasticity of the skin decreases, and those expressions that occur frequently leave their traces; facial lines, furrows, and pouches become etched into the surface as relatively permanent features.

Asymmetry

Faces are structurally asymmetric, often with one side larger than the other. Structural asymmetry, approximated by distance from facial landmarks to center points, ranges from 4 to 12% average difference, depending on the landmark measured [14]. The right side tends to be larger, and facial landmarks on the right side tend to be rotated more inferiorly and posterior to than those on the left [14]. Facial

asymmetry is perceptually salient (Fig. 5) and can result from multiple factors. These include genetic variation, growth, injury, age, and depending on type of asymmetry, sex.

Recent evidence suggests that individual differences in asymmetry may be a useful biometric. When asymmetry metrics were added to a baseline face recognition algorithm, Fisher-Faces, recognition error in the FERET database decreased by close to 40% [15]. These findings are for 2D images. Because some aspects of asymmetry are revealed only with 3D measurement, error reduction may be greater when 3D scans are available.

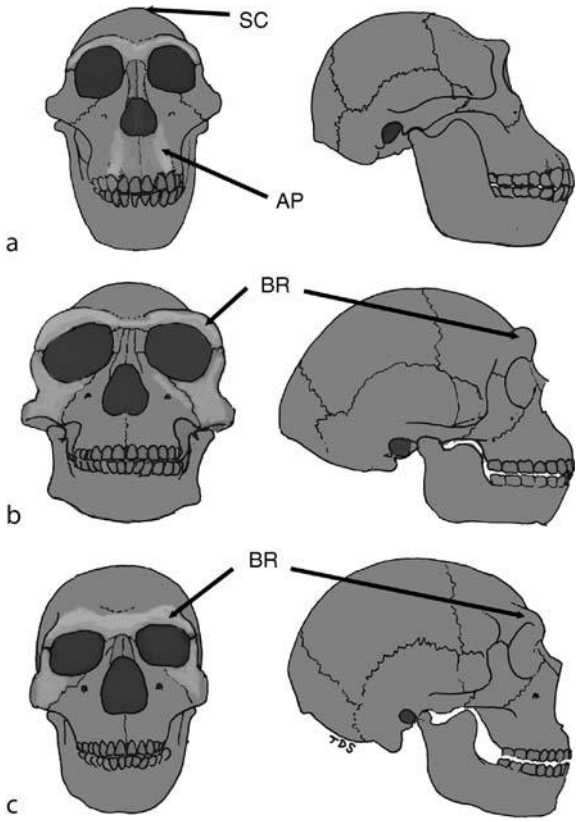
Another factor that may contribute to the appearance of asymmetry is facial expression. While most of the variation in asymmetry at peak expression is accounted for by structural asymmetry (i.e., basal or intrinsic asymmetry at rest) [16], movement asymmetry contributes less but significant variance to total asymmetry. A function of movement asymmetry may be to attenuate or exaggerate apparent asymmetry. The influence of facial expression in face recognition has been relatively little studied.

Evolution of Human Face Forms

The first recognizable human ancestor was *Australopithecus*. The gracile (slender or slight) australopithecines, such as *A. africanus*, are direct ancestors to *Homo* and modern humans. The craniofacial skeleton of the gracile australopithecines is characterized by having relatively large brains when compared to modern apes (but smaller than *Homo*) and massive molar teeth with large jaws. Large jaws need large muscles to move them, which in turn leave large muscle markings such as the sagittal crest and a flaring mandibular angle. Powerful chewing stresses were dealt with in the facial skeleton by placing anterior pillars on either side of the nasal apertures. These anterior pillars were massive vertical columns supporting the anterior part of the hard palate. Any facial mask stretched over this facial skeleton would have been influenced in appearance by these bony features. Overall, australopithecines had a dolicocephalic head with a prominent, prognathic "snout" relative to modern humans [17].

In *Homo erectus*, the "snout" is greatly reduced as are the molars (Fig. 6). The sagittal crest and anterior

Anatomy of Face. **Figure 5** Left: original face images taken under balanced bilateral lighting. Middle: a perfectly symmetrical face made of the left half of the original face. Right: a perfectly symmetrical face made of the right half of the original face. Notice the difference in nasal regions in both individuals caused by left–right asymmetry of the nasal bridge. [14]. © Elsevier.



Anatomy of Face. **Figure 6** Frontal (left) and right views (right) of fossil humans. a.) *Australopithecus africanus*, b.) *Homo erectus*, c.) *H. neanderthalensis*. Abbreviations: AP: anterior pillar; SC: sagittal crest; BR: brow ridges. Note the relatively small neurocranium in *A. africanus* and the relative states of dolicocephaly and leptoproscopy, reflecting the small brain. Note also the anterior pillars and massive jaws. While a brow ridge is present in this species, it is relatively small compared to *Homo*. In *H. erectus*, note the enlarging neurocranium and wider face with a reduced "snout", reflective of the enlarging brain in this species relative to *A. africanus*. Additionally, the anterior pillars have disappeared and the size of the jaw is reduced but the brow ridges enlarge. Similarly, *H. neanderthalensis* has an even larger brain and greater reduction of the "snout" relative to *H. erectus*. © Tim Smith

pillars thus disappear and the head shape becomes more brachycephalic as in modern humans, due to the dramatic increase in brain size. The nasal aperture becomes much wider, and the nares in this species attain the downward facing posture as in modern humans. A prominent brow ridge develops in *H. erectus* that is lost in modern humans [17].

Neanderthals, *H. neanderthalensis*, are the most recent fossil human. Their brain size was actually larger than that of modern humans. Neanderthals are generally characterized by an enormous nasal opening, a reduced snout relative to *H. erectus* but larger than in modern humans, and a swollen, "puffy" appearance to the face in the region of the malar bones [17].

What might the face have looked like in each of these fossil humans? What might their facial expression repertoire have been? Facial musculature does not leave muscle markings behind on the bones so it cannot be described with any degree of certainty. However, since the mimetic musculature in primates follows a very conservative pattern from the most primitive strepsirhines through humans [8], it is logical to assume that mimetic musculature in fossil humans was very similar to our own and to chimpanzees, our closest living relative.

Conclusions

Variation in facial appearance among human individuals is considerable. While the mimetic musculature produces facial movements, of which facial expressions of emotions are best known, it is not the major source of this variation. The major source is in the facial skeleton itself. Three representative head types have been identified, dolico-, meso-, and brachycephalic. These types correspond to geographic dispersion of human populations over the past 50,000 years or more. Within each of these types, there is considerable variation, which is likely to increase in light of demographic trends. Such individual differences in facial anatomy have been relatively neglected in face recognition research. Asymmetry is a recent exception. Preliminary work in 2D images suggests that inclusion of asymmetry metrics in algorithms may significantly reduce recognition error. Because many asymmetry metrics are 3D, their relative utility may be even greater where 3D imaging is feasible. Asymmetry, of course, is only one type of individual variation in facial anatomy. Others are yet to be explored. The anatomical record suggests that such work could be promising.

Fossil humans had facial skeletons drastically different from contemporary humans, *Homo sapiens*. In general, human facial skeletons have evolved from a long, narrow form with a prominent "snout" to one in which the face is more "tucked under" the braincase.

Facial expression and face recognition are major components of communication among humans. Understanding the evolution of the human facial form provides a window for an understanding of how and why so much emphasis is placed on the face in recognition of individual identity.

Acknowledgements

Preparation of this manuscript was supported in part by grant NIMH R01–501435 to the University of Pittsburgh. The authors wish to thank Bridget M. Waller for much thoughtful discussion on the topic of muscle variation in humans and helpful comments on earlier versions of this work. Fig. 1, 4, and 6 by Timothy D. Smith.

Related Entries

► Eye Features and Anatomy
► Face, Forensic Evidence of
► Face Recognition, 3D
► Face Variation

References

1. Phillips, P.J., Scruggs, W.T., O'Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., et al.: FRVT 2006 and ICE 2006 large-scale results. Technical Report NISTIR 7408. National Institute of Standards and Technology Washington: (2007)
2. Enlow, D.H., Gans, M.G.: 1Essentials of facial growth. W. B. Saunders, Philadelphia (1996)
3. Schwartz, J.H.: Skeleton keys: An introduction to skeletal morphology, development, and analysis (2nd edn). Oxford University Press, New York (2007)
4. Enlow, D.H.: Facial growth (3rd edn). W.B. Saunders, Philadelphia (1990)
5. Mooney, M.P., Siegel, M.I.: Overview and introduction. In: Mooney M.P., Siegel M.I. (eds.) Understanding craniofacial anomalies: The Etiopathogenesis of Craniosynostoses and Facial Clefting, pp. 3–10. Wiley-Liss, New York (2002)
6. Pessa, J.E., Zadoo, V.P., Adrian, E.J., Yuan, C.H., Aydelotte, J., Garza, J.R.: Variability of the midfacial muscles: Analysis of 50 hemifacial cadaver dissections. Plast. Reconstr. Surg. **102**, 1888–1893 (1998)
7. Waller, B.M., Cray, J.J., Burrows, A.M.: Facial muscles show individual variation only when non-essential for universal facial expression. Emotion. **8**, 435–439 (2008)
8. Burrows, A.: Primate facial expression musculature: Evolutionary morphology and ecological considerations. Bio Essays. **30**, 212–225 (2008)
9. Pessa, J.E., Zadoo, V.P., Garza, P., Adrian, E., Dewitt, A., Garza, J.R.: Double or bifid zygomaticus major muscle: Anatomy, incidence, and clinical correlation. Clin. Anat. **11**, 310–313 (1998)
10. Sato, S.: Statistical studies on the exceptional muscles of the Kyushu - Japanese Part 1: The muscles of the head (the facial muscles). Kurume Med. J. **15**, 69–82 (1968)
11. Goodmurphy, C., Ovalle, W.: Morphological study of two human facial muscles: orbicularis occuli and corrugator supercilii. Clin. Anat. **12**, 1–11 (1999)
12. Schmidt, K.L., Cohn, J.F.: Human facial expressions as adaptations: Evolutionary perspectives in facial expression research. Yolk. Phys. Anthropol. **116**, 8–24 (2001)
13. Larrabee, J.W.F., Makielski, K.H.: Surgical anatomy of the face. Raven Press, New York (1993)
14. Ferrario, V.F., Sforza, C., Ciusa, V., Dellavia, C., Tartaglia, G.M.: The effect of sex and age on facial asymmetry in healthy participants: A cross-sectional study from adolescence to mid-adulthood. J. Oral Maxillofac. Surg. **59**, 382–388 (2001)
15. Liu, Y., Schmidt, K.L., Cohn, J.F., Mitra, S.: Facial asymmetry quantification for expression invariant human identification. Comput. Vision Image Underst. **91**, 138–159 (2003)
16. Schmidt, K.L., Lui, Y., Cohn, J.F.: The role of structural facial asymmetry in asymmetry of peak facial expressions. Laterality **11**(6), 540–561 (2006)
17. Tattersall, I., Schwartz, J.H.: Extinct humans. Westview, Boulder, CO (2000)

Anatomy of Fingerprint

► Anatomy of Friction Ridge Skin

Anatomy of Friction Ridge Skin

R. Austin Hicklin
Noblis, Fairview Park Drive, Falls Church, VA, USA

Synonyms

Anatomy of Fingerprint; Palmprint anatomy

Definition

Friction ridge skin refers to the skin of the palms of the hands and fingers as well as the soles of the feet and toes.

Friction ridge skin can be differentiated from the skin of the rest of the body by the presence of raised ridges, by epidermis that is thicker and structurally more complex, by increased sensory abilities, by the absence of hair, and by the absence of sebaceous glands. The presence of friction ridges enhances friction for skin used in grasping. Note that the term ▶ *fingerprint* refers to an impression left by the friction skin of a finger rather than the anatomical structure itself.

Introduction

The palms of the hands and fingers as well as the soles of the feet and toes have skin that is distinctly different from the skin of the rest of the body. This skin is known as thick skin, volar skin, or hairless skin by anatomists, but is known as friction ridge skin in the biometric and forensic communities due to the distinctive patterns of raised ridges that can be used in identification.

Surface Features

Friction ridge skin is covered with a corrugated texture of ridges that enhance the ability of the hand (and feet) to grasp or grip surfaces. The ridges are three-dimensional structures with irregular surfaces, separated by narrower furrows or valleys. The surface features of friction ridge skin are often divided into three levels of detail: ▶ ridge flow and pattern for an area of skin (level-1 features); ridge path and ▶ minutiae for a specific ridge (level-2 features); and dimensional, edge shape, and pore details within a specific ridge (level-3 features) [1, 2].

The morphological patterns of ridge flow vary with the location. When comparing the areas of friction ridge skin, the most complex patterns can usually be found on the outermost (distal) segments of the fingers, at the interdigital portion of the palm across the bases of the fingers, on the tips of the toes, and at the portion of the sole across the bases of the toes. The ridges in these areas often have tightly curving patterns with continuously changing direction. The complexity of ridge flow in these areas is because of the fetal development of volar pads in those areas (discussed below in Friction Skin Development). The other areas of friction skin, such as the extreme tips and lower joints of the fingers, and the lower portion of the palm, usually contain gently curving ridges without dramatic changes in direction.

For the distal (outermost) segments of the fingers, ridge flow falls into three general pattern classifications: (1) whorls, in which the ridge flow forms a complete circuit; (2) loops, in which the ridge flow enters from one side, curves, and returns in the same direction from which it came; and (3) arches, in which the ridge flow enters from one side and exits the opposite side. The most common patterns are ulnar loops, or loops in which the flow points to the ulna (the bone in the forearm closest to the little finger). The most complex patterns (double loop, central; pocket loop, and accidental) are considered subclasses of whorls. In very rare circumstances, friction skin is composed of dissociated small sections of ridges that do not form continuous ridges, a genetic condition known as dysplasia [1].

Ridges are of varying lengths, and may be as short as a segment containing a single pore, or may continue unbroken across the entire area of friction skin. The points where specific ridges end or join are known as minutiae, and are of particular interest: ridge endings and bifurcations are the features most frequently used in identification. Very short ridges containing a single pore are known as dots. Many fingerprints have thin, immature, often discontinuous ridges known as incipient ridges between the primary ridges as shown in Fig. 2.

The ridges vary markedly in diameter and frequency between different parts of the body: for example, the ridges of the soles of the feet are notably coarser than those of the palms and fingers, and the ridges of the little fingers are often finer than those of the other fingers. The diameter of ridges increases with an individual's size, with male ridges generally larger than for females, and adult ridges notably larger than for children. Within a given small section of skin, some ridges may be finer or coarser than the others in the surrounding area.

The ridges are punctuated by a series of sweat pores. While on average the spacing of the pores is relatively regular, the specific locations of pores are distinctive features that are used in identification.

Friction skin flexes along lines known as flexion creases. The most prominent of the flexion creases are the interphalangeal creases that separate the segments of the fingers, and the thenar and transverse creases of the palm. A variety of minor flexion creases are particularly notable on the palm. Flexion creases form along areas in which the skin is more strongly attached to the underlying fascia [3]. The smallest of the flexion creases are known as white lines, which occur randomly over the skin [1]. The prevalence and depth of white lines

increases with age. White lines are especially prevalent on the lower joints of the fingers, and on the thenar (base of the thumb). In some cases, a large number of white lines make the underlying ridges difficult to discern, as shown in Fig. 1.

Friction Skin Structure

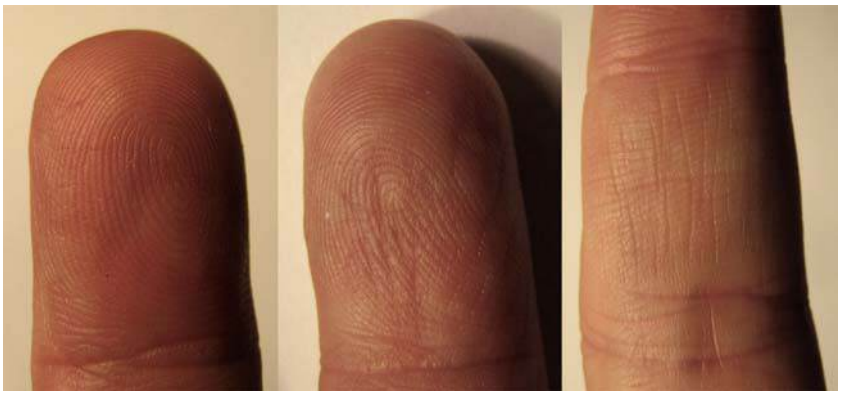
Skin is a protective barrier that contains nerve receptors for a variety of sensations, regulates temperature, allows the passage of sweat and sebaceous oils, and houses the hair and nails. Friction ridge skin is differentiated from thin skin not just by the presence of raised papillary ridges, but also by epidermis that is much thicker and structurally more complex, by increased sensory abilities, by the absence of hair, and by the absence of sebaceous glands.

Skin throughout the body is composed of three basic layers: the hypodermis, dermis, and epidermis.

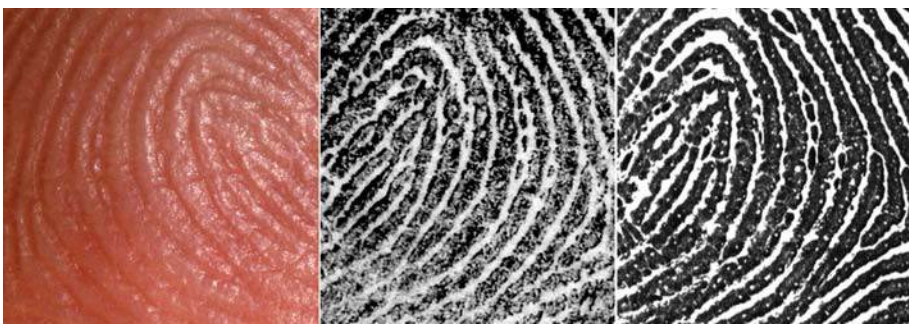
The innermost hypodermis (also known as subcutaneous tissue or the superficial fascia) is made of connective tissue that stores fat, providing insulation and padding. The hypodermis varies in thickness, but is particularly thick in friction ridge skin [3, 4].

The dermis is composed of dense connective tissue that provides strength and plasticity. The dermis houses blood vessels, nerve fibers and endings, and sweat glands. In non-friction ridge skin, the dermis also contains sebaceous (oil) glands, hair follicles, and arrector pili muscles, which raise hair and cause "goose bumps" [4].

The boundary between the dermis and epidermis is of particular interest for friction ridge skin. The dermis and epidermis are joined by papillae, which are columnar protrusions from the dermis into the epidermis, and rete ridges, which are the areas of the epidermis



Anatomy of Friction Ridge Skin. **Figure 1** Friction skin ridge flow: whorl and loop finger patterns, and unpatterned skin from a lower finger joint. Note the minor creases (white lines), especially in the middle and right images.



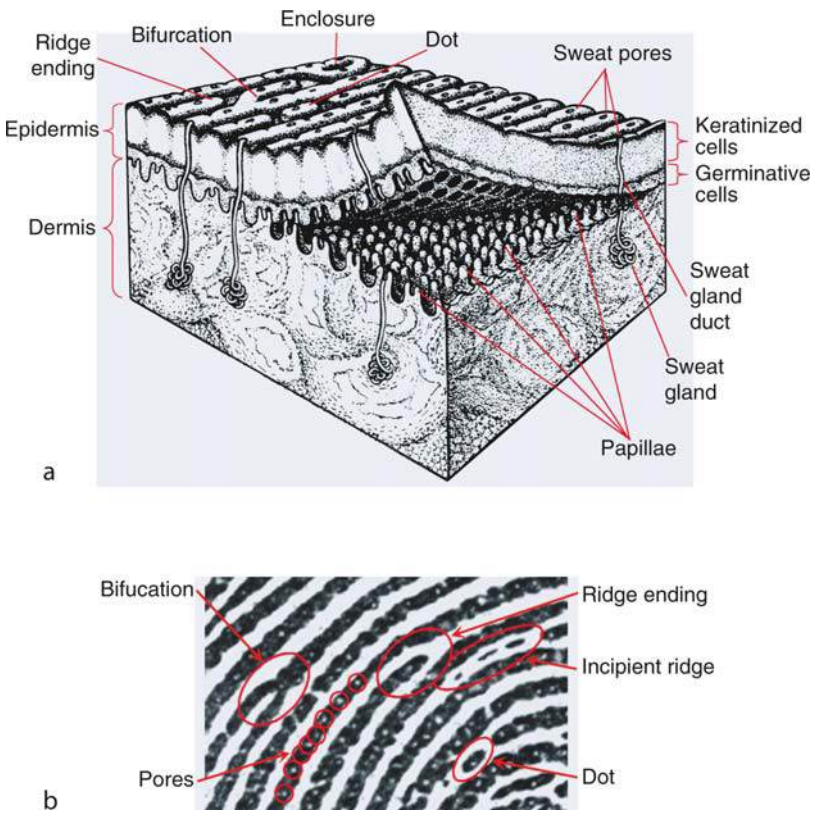
Anatomy of Friction Ridge Skin. **Figure 2** Friction ridge skin with corresponding inked and optical live scan fingerprint impressions. Note the variation in appearance of details, especially the incipient ridges. The pores are clearly visible in the rightmost image.

surrounding the papillae. The papillae anchor the epidermis and contain nerve endings and blood vessels. Papillae in thin skin are small, relatively infrequent, and are not arranged in any particular pattern. In friction ridge skin, the papillae are densely arranged in curved parallel lines by pairs, with pairs of papillae surrounding the sweat pores [3] (Fig. 3).

The epidermis provides the outermost protective layer, and is structurally very different between friction ridge and thin skin. The epidermis does not contain blood vessels, and therefore the basement membrane, which joins the dermis and epidermis, serves as the medium through which nutrient and waste passes. The lowest level of the epidermis (stratum basale) contains a single layer of basal generative cells, which are anchored to the basement membrane. These basal cells continuously create columns of new cells. It takes 12–14 days for a cell to progress from the innermost basal layer to the outermost horny or cornified layer of the epidermis. During this time, the cell flattens forms interconnections with the neighboring cells, is filled with keratin (the tough protein found in hair and nails), and dies. The dead cells are continuously exfoliated, with the entire epidermis being renewed approximately every 27 days. The thickest portion of the cornified layer of cells is generated along the lines of paired papillae, resulting in visible friction ridges, punctuated with pore openings. The epidermis in friction ridge skin is 0.8–1.44 mm thick, as compared to 0.07–1.12 mm thickness elsewhere. Heavy use can result in substantially increased thickness of the epidermis, in the form of calluses or corns [3, 4].

Friction Skin Development

The individual characteristics of friction ridge skin are determined during fetal development, based on a combination of genetic and random factors. The overall pattern of friction ridges is determined by the formation and regression of volar pads in the fetus. Starting at approximately 6 or 7 weeks of gestational age, human fetuses form swellings of tissue in what will later become the dermis: 11 of these volar pads generally



Anatomy of Friction Ridge Skin. **Figure 3** (a) Structure of friction ridge [5]. (b) Examples of friction ridge features.

develop on each hand, with 1 at each fingertip, 4 interdigital pads at the bases of the fingers, 1 thenar pad at the ball of the thumb, and 1 hypothenar pad along the outside of the palm. Each foot has 11 pads in corresponding locations. The size, shape, and period of development of the volar pads are determined to a large extent by genetics. The pads continue to grow for a few weeks and then regress as the rest of the hands and feet grow. The volar pads are usually no longer evident by about 16 weeks of gestational age. During the period of volar growth and regression, starting at about 10 weeks of gestational age, the basal epidermal cells begin a stage of rapid proliferation, especially surrounding the sweat glands. Since this process occurs while the volar pads are regressing, the result is that the growing cells fuse together along the lines of stress created by the collapse of the volar pads. While the overall form of the ridges follows the contours of volar pads, the specific paths, bifurcations, and endings of the ridges are determined by the stresses encountered during growth [1, 6, 7].

The overall form of the ridges is determined by the topography of the volar pads, with the pattern class determined by the height and symmetry of the volar pads. This can be seen most easily in examining the areas without volar pads, such as the lower joints of the fingers and the lower palm: the ridge flow in these areas is generally simple, with ridges flowing across the area without dramatic changes in direction. If volar pads are small the resulting pattern will be an arch, with simple ridge flow similar to the areas without volar pads. If the volar pads are large and centered, the resulting pattern will be a whorl, with ridge flow following the circuit of the pad.

Because of the genetic basis for volar pad formation, overall ridge flow or pattern classification is often similar between siblings, especially identical twins. For the same reason, fingerprint patterns on an individual's left and right hands are often similar to mirror images of each other. However, because the path of any individual ridge results from chaotic stresses, the details of minutiae are specific to the individual.

Prevalence

Friction skin covers the palms and soles of all anthropoid primates (monkeys, apes, and humans), as well as on portions of the prehensile tails of some New World monkeys. Some but not all prosimian primates (lemurs) have friction skin on portions of their palms and soles [8]. Friction skin is unusual in other mammals, but is found on portions of the grasping hands and feet of two species of tree-climbing marsupials (koalas and one form of phalanger) [9]. Note in all cases that friction ridge skin is associated with grasping surfaces: the ridges increase friction, and the greater density of nerve endings improves tactile sensitivity.

Problems in Capturing Friction Skin Features

Friction ridge skin is a flexible, three-dimensional surface that will leave different impressions depending on factors including downward or lateral pressure, twisting, and the medium used. Even when only considering clear impressions, the details of fingerprints and ► palmprints vary subtly or substantially between impressions. As downward pressure increases, the apparent diameter of the valleys decreases and the ridges widen. The frequency of ridges is affected by lateral compression or stretching. A bifurcation of a physical ridge does not always appear as a bifurcation in the corresponding print, but may appear to be a ridge ending under light pressure. Incipient ridges may become more discontinuous or vanish altogether under light pressure. Pores are not always evident in fingerprints even at high resolution, which can be explained in part by the tendency to fill with liquid such as sweat or ink. This variability between different impressions of an area of friction skin is responsible for much of the complexity of matching fingerprints, whether performed by human experts or automated recognition systems.

Related Entries

► Anatomy of Hand
► Fingerprint Classification
► Fingerprint Individuality
► Fingerprint Recognition Overview
► Palmprint Feature
► Palmprint Matching

References

1. Ashbaugh, D.R.: Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advance Ridgeology. CRC Press, Boca Raton, Florida (1999)
2. Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST): Glossary, Version 1.0. http://www.swgfast.org/Glossary_Consolidated_ver_1.pdf (2003)
3. Standring, S. (ed.): Gray's Anatomy: The Anatomical Basis of Clinical Practice, 39th edn. Elsevier, London (2004)
4. Weiss, L. (ed.): Cell and Tissue Biology: A Textbook of Histology, 6th Edition. Urban & Schwarzenberg, Baltimore (1988)
5. Federal Bureau of Investigation: The Science of Fingerprints, Rev 12–84. U.S. Government Printing Office, Washington, DC (1984)
6. Maceo, A.: Biological Basis of Uniqueness, Persistence, and Pattern Formation. In: 4th International fingerprint symposium, Lyon, France, 17–19 May 2006. http://www.interpol.int/Public/Forensic/fingerprints/Conference/May2006/presentations/2AliceMaceoPart1.pdf, 2AliceMaceoPart2.pdf (2006)
7. Wertheim, K., Maceo, A.: The Critical Stage of Friction Ridge and Pattern Formation. J. Forensic Ident. **52**(1), 35–85 (2002)
8. Ankel-Simons, F.: Primate Anatomy: An Introduction. Academic Press, San Diego (2000)
9. Henneberg, M., Lambert, K.M., Leigh, C.M.: Fingerprint homoplasy: koalas and humans. naturalSCIENCE.com. Heron Publishing, Victoria, Canada (1997)

Anatomy of Hand

Amioy Kumar[1], Tanvir Singh Mundra[2], Ajay Kumar[3]

[1,2]Biometrics Research Laboratory, Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India

[3]Department of Computing, The Hong Kong Polytechnic University

Synonyms

Hand physiology; Hand structure

Definition

The ▶ anatomy of human hand is quite unique and includes the configuration of bones, joints, veins, and muscles. The physiological interconnection and structure of these parts are responsible for the structure of the human hand. The functional area of hand includes the five fingers, palm, and the wrist. Among a number of biometric modalities that are used for human identification, hand-based modalities achieve high performance and have very high user acceptance. A hand-based biometric system integrates several physiological and/or behavioral features that have their individuality in the anatomy of hand. The prime focus of this study is on internal and physiological structure of human hand which defines the uniqueness of various hand related biometric modalities.

Introduction

The anatomical study of human hand is not new; it dates back to prehistoric times, but it is finding new applications in the field of biometrics. The proper understanding of structure requires the knowledge of function in the living organism. As one of the basic life sciences, anatomy is closely related to medicine and to other branches of biology. The hands of the human being are the two multi-fingered body parts located at the end of each arm. It consists of a broad palm with five fingers, each attached to the joint called the wrist. The back of the hand is formally called the dorsum of the hand. The uniqueness of the human hand, as compared to the other animals comes from the fact that all the fingers are independent of each other and the thumb can make contact with each finger.

The anatomy of hand is the key to ascertain the individuality of hand-based biometrics. The hand-geometry biometric largely represets the anatomy of hand bones and muscles. The hand-vein biometric represents the uniqueness in the anatomy of hand-veins while the palmprint represents ▶ epidermis on the palm. The behavioral biometrics like signature is also highly dependent on the anatomy of bones and muscles. Therefore the study of hand anatomy is fundamental to ascertain the individuality of hand-based biometrics.

Structure of the Human Hand

The internal structure of hand is an assortment of bones, muscles, nerves, and veins.

Bones

The structure of the hand is primarily attributed to the bones comprising the human hand. The hand is composed of 27 bones, broadly divided into three groups called carpals, metacarpals, and phalanges (Fig. 1). The wrist of the hand consists of a cluster of bones named as carpals. These bones are considered as a part of wrist and are responsible for the to-fro and back-forth movement of the wrist. These are eight in number and are named as:

1.	Scaphoid	2.	Lunate
3.	Triquetrum	4.	Pisiform
5.	Trapezium	6.	Trapezoid
7.	Capitate	8.	Hamate

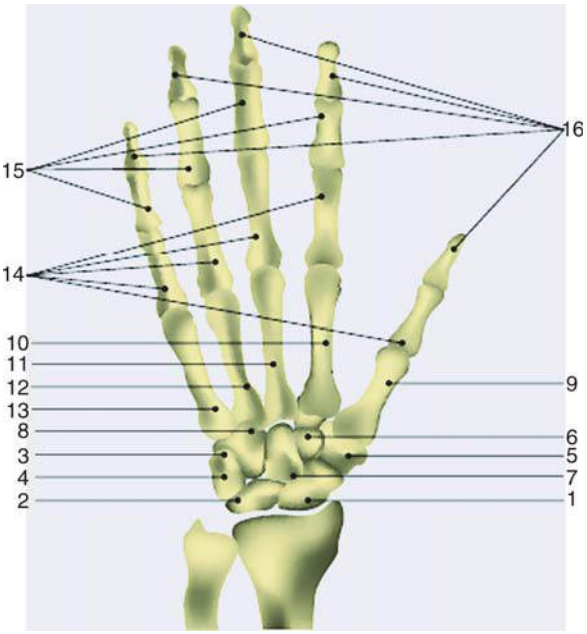
Metacarpals are the intermediate part of the fingers and the wrist [1]. This cluster of bones make the central part of the hand called the palm. The metacarpals are five in number and are named as:

9.	First metacarpal (Thumb)	10.	Second metacarpal (Index finger)
11.	Third metacarpal (Middle finger)	12.	Fourth metacarpal (Ring finger)
13.	Fifth metacarpal (Little finger)		

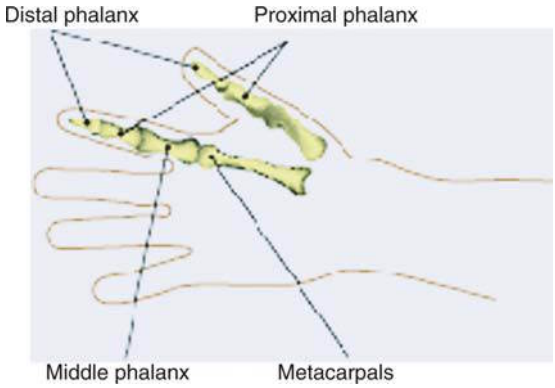
The remaining fourteen bones are called the phalanges. These are named as follows:

14.	Proximal	15.	Medial
16.	Distal		

There are two in the thumb, and three in each of the four fingers, as shown in (Fig. 2). The distal phalanges carry the nails, the middle phalanges are in the middle and the proximal phalanges are closest to the palm. Bones are the most important structure of the human hand and responsible for almost all the activities of the hand. Even so, bones structures in the hand are not a popular candidate for biometric authentication. Being a hidden structure of hand, the acquisition of hand bone images is very difficult. However, the hand bone structures are useful in forensic identification especially in situations when other physiological structures are not available, e.g. during accidents/fire. The individuality of hand bone structures is generally believed to be low due to the high similarity in the bone types.



Anatomy of Hand. **Figure 1** Skeletal structure of the human hand.



Anatomy of Hand. **Figure 2** Phalanx bones of the human hand.

Muscles

Muscles are like the building blocks on the bones. These not only make the hand robust in gripping but also are very helpful in its movement. The muscles of the human hand are composed of two types of tissue, namely the extrinsic muscle groups and intrinsic muscle groups [3]. The extrinsic groups of muscles are generally present in dorsal (back) part of the hand, or palmer (grasping) part of the hand. It is broadly

divided into extensors, present on dorsal part and flexors, present on the palmar part of the hand. The extensor muscles are further divided into those whose movement is around wrist as:

1.	Extensor carpi radialis longus	2.	Extensor carpi radialis brevis
3.	Extensor carpi ulnaris		

4.	Abductor pollicus longus	5.	Extensor pollicus brevis
6.	Extensor pollicus longus	7.	Extensor digiti minimi
8.	Extensor digitorium		

And those whose movement is around digits (the four fingers without the thumb) of hand as:

All the extensor muscles are shown in (Fig. 3). Unlike extrinsic muscles, the intrinsic muscles of the hand are originated at wrist and hand. It can be divided as: Dorsal and Volar muscles. The dorsal intrinsic muscles (Fig. 4) can be further subdivided into:

1.	Dorsal interossei	2.	Abductor digiti minimi

The volar intrinsic muscles are present in two layers:

1. Superficial layer

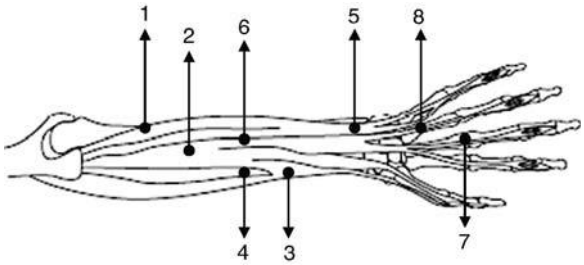
1.1.	Abductor digiti minimi	1.2.	Flexor digiti minimi
1.3.	Lumbricals	1.4.	Adductor pollicis
1.5.	Abductor pollicis brevis		

2. Deep layer

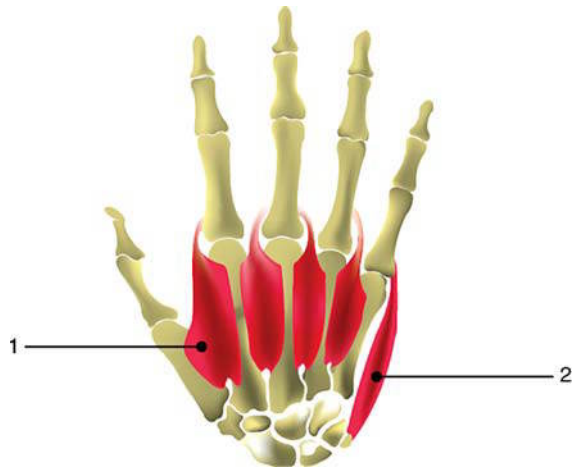
2.1.	Oppones digiti minimi	2.2.	Palmar interossei

The superficial and deep muscles are shown in Figs. 5a and 5b.

Due to intrinsic features of human hand, the muscles are weak candidate for biometric identification. Muscles are covered by skin and hence it is very difficult for an imaging system to capture muscle structures independently. The acquisition of muscle structure requires very complex imaging techniques, such as magnetic resonance imaging which is very expensive. While capturing hand geometry or palmprint images, the muscles have very little or no effect on hand surface when the user-pegs are employed to constrain the hand movement. However, the



Anatomy of Hand. **Figure 3** Extrensor muscles of the hand.

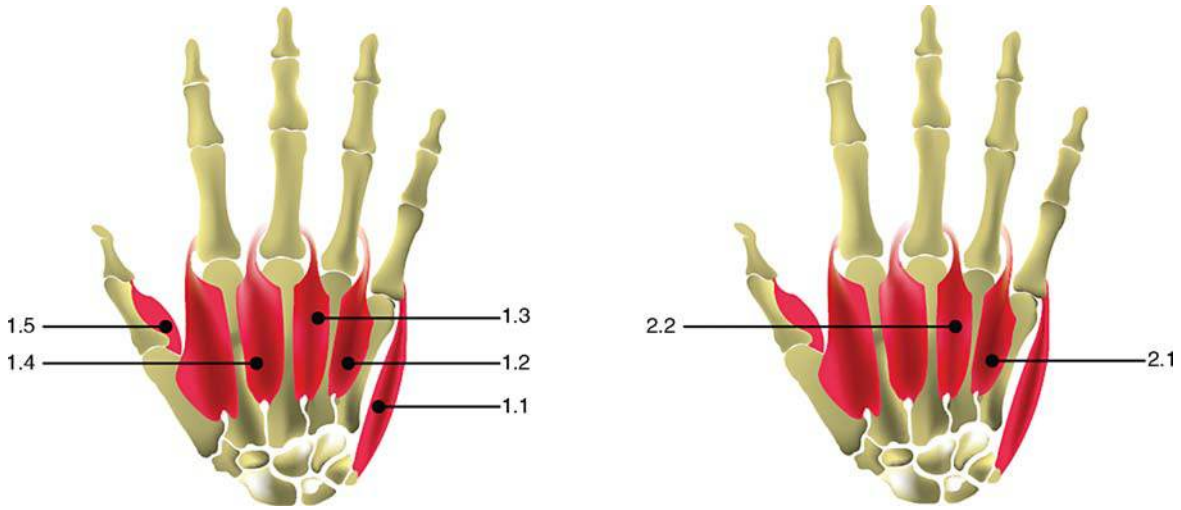


Anatomy of Hand. **Figure 4** Dorsal intrinsic muscles.

peg-free hand imaging introduces some effect due to the independent movement of fingers. One of the major weaknesses with muscles as biometric trait is that with the change in age, it begins to loose its shape and strength. Due to change in shape the hand surface of a young man looks quite different as compared to an old man. However, being an internal part of hand surface, muscles are quite stable with respect to changes in humidity and temperature. Another advantage with the possible usage of muscle as a biometric trait is that being hidden structure it is very difficult to spoof and any change in muscle structure requires very complex surgical operations.

Nerves

The nerves are a very important part of hand and helpful in sensing objects. These internal structures are also responsible for carrying the sensory information from one part of the body to the other.



Anatomy of Hand. Figure 5 (**a**) Superficial layers. (**b**) Deep layer.

These nerves are quite stable and unique candidate for potential biometric identification. There are two ways of discussing nerve distribution in wrist and hand of human body. These are:

The peripheral nerves are distributed around wrist and hand and can be classified as:

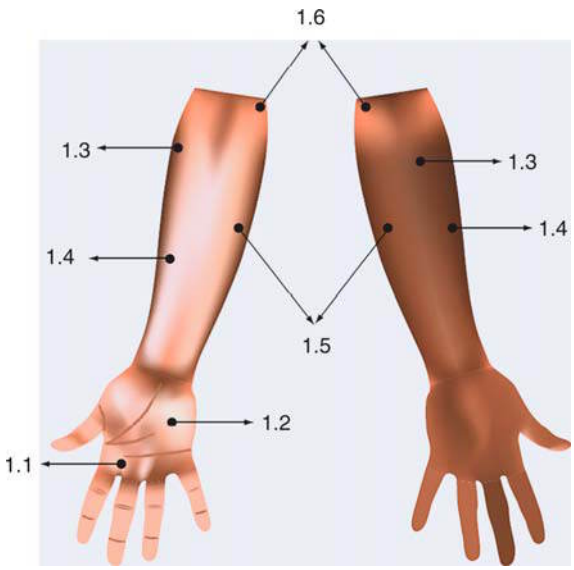
1.	Peripheral nerves or	2.	Dermatomes nerves

The dermatomic regions of the skin are very sensitive from medical point of view, as pain in this area indicates spinal damage. Nerves in these areas are originated from dorsal root (single spinal nerve root) [4]. These root nerves are:

1.1.	Median nerve	1.2.	Ulnar nerve
1.3.	Radial nerves	1.4.	Lateral cutaneous nerve of forearm
1.5.	Medial cutaneous nerve of forearm		(musculocutaneous nerve)

1.1.	C5	1.2.	C6
1.3.	C7	1.4.	C8
1.5.	T1	1.6.	T2

Nerve root C5 is associated with radial nerve, C6 is associated with median nerve, and C7 is associated with both median and radial nerve. C8 forms the median, ulnar, and radial nerve. T1 root is of Medial cutaneous nerve of forearm. All the root and peripheral nerves are shown in Fig. 6.



Anatomy of Hand. Figure 6 The peripheral nerves distribution in arm and wrist of human hand.

However, nerves are also the hidden structure and therefore very difficult to be imaged. This is the principal reason why the nerve structures are not yet been explored for biometric identification.

Palmprint

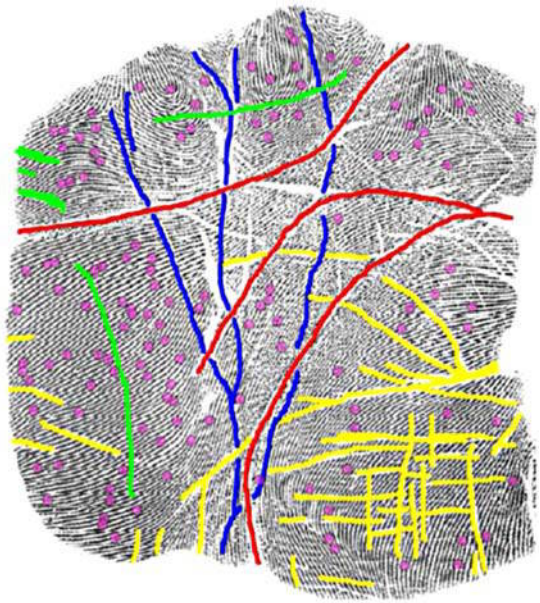
The human *palm* is defined as the inner portion of the hand starting from the wrist to the root of the fingers.

The *print* is an impression made when the body part is pressed against some surface. A palmprint therefore illustrates the physical properties of skin pattern such as lines, points, minutiae, and texture [2]. Palmprint identification can be seen as the capability to uniquely identify a person amongst others, by an appropriate algorithm using the palmprint features. The palmprint features are mainly developed during the life processes, due to biological phenomenon, with the growth of fetus in the uterus. Even a minute change in these inherent phenomenon, changes the complete life process and hence the structure of two different palms is expected to be never the same. The main features of interest from the palmprints are as follows:

1. Minutiae features from the palm friction ridges
2. Principal lines, which are the most, darken lines on the palm
3. The thinner and irregular lines, as compared to principle lines called wrinkles
4. Datum points, which are the end points of principal lines

In the most palmprint recognition approaches, various texture features are acquired from the 2D images. However, a limitation of such a system is that the acquired images and hence the accuracy of such systems is highly sensitive to the illumination changes. Recent research in this area has shown promising results using simultaneously acquired 3D palmprint features [5]. A 3D scanner can be used to capture the palmprint surface. Such acquisition not only reduces the effect of illumination, but also provides a better curvature of principle lines, depth, and wrinkles of the palmprint. The palmprint systems employing 3D palmar features are certainly more reliable and robust to security threats as compared to those systems employing only 2D features.

Most of the above discussed palmprint features are acquired from low resolution images (approximately 100 pixels per inch). Such extracted features and matching algorithms cannot suit a typical forensic application. More palmprint features such as: palmar friction ridges, palmar flexion creases, palmar texture, minutiae etc., can be utilized for recognition purposes. Friction ridges are folded pattern of palm skin with sudoriferous gland but without hair. The palmar friction ridges are formed during the embryonic skin development but after the appearance of flexion creases [6, 7]. The palmer friction ridges originate from the deeper ▶ dermis layer within the first twelve weeks of fetal development.



Anatomy of Hand. **Figure 7** The palmar flexion creases (adopted from [6]): major creases (red), minor

As shown in Fig. 7, the flexion creases appearing on palmar surface can be grouped in three categories: major flexion, minor flexion, and secondary creases. The major flexion creases are the largest creases and include distal transverse (heart-line), radial transverse (life-line), and proximal transverse (head-line). These major flexion creases are highly visible large lines that are often employed as reference while aligning two palmprints for biometric identification. The minor flexion creases, along with the secondary creases and minutiae locations, serve as reliable features for palmprint identification for forensic [6] and civilian applications.

Fingerprint

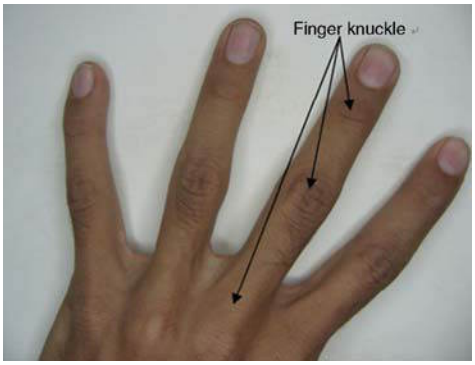
The impression of friction ridges formed from the inner surface of fingers and thumb is referred to as fingerprint. The formation of finger tips is similar to the formation of blood vessels or capillaries during the growth of fetus in the uterus. The formation of skin and the volar surface of palm or sole in the fetus are due to the flow of amniotic fluids in a micro-environment. With the minor change in the

flow of amniotic fluids and the position of fetus in the uterus, the minute skin structures around palm or finger tips begins to differentiate. Thus, the finer details present on fingertips are determined by very minute biological phenomenon in a micro-environment. Even a small difference in micro-environment, changes the process of cell formation completely and these structures vary from hand to hand. The similarity of these minute structural variations is virtually impossible to detect [8]. In biometrics literature, fingerprint is treated as one of the most reliable modality due to its high structural variance from hand to hand as even identical twins have different fingerprints [9]. A fingerprint broadly consists of a pattern of ridges/valleys. Its uniqueness is attributed to the ridge characteristics and their inter-relationship. Minutiae points in the fingerprints are defined as ridge endings, the point where the ridges end abruptly, or ridge bifurcation/trifurcation, where the ridges are divided into different branches. These patterns have shown to be quite immune to aging and biological changes. The fingerprint features are extracted from the characteristics of frictin ridges and generally into three categories:

1. Macroscopic ridge ridge flow patterns (core and delta points)
2. Minutiae features (ridge endings and bifurcations)
3. Pores and ridge contour attributes (incipient ridges, pore, shape and width)

Finger Knuckle

The joints from phalanx bones (Fig. 2) of human hands generate distinct texture patterns on the finger-back surface, also known as the dorsum of hand. In particular, the image pattern formation from the finger-knuckle bending is quite unique and makes this surface a distinctive biometric identifier. Figure 8 identifies three finger knuckles, from each of the finger, which can be potentially employed for personal identification. The anatomy of fingers allows these knuckles to bend forward and resist backward motion [2]. Therefore each of three finger-knuckle (Fig. 8) results in a very limited amount of crease and wrinkles on the palm-side of the fingers. The anatomy of joints from the phalanax bones results in a greater amount of texture pattern from the middle finger knuckle surface, from each of the fingers, and has emerged as another promising modality for the biometric identification.



Anatomy of Hand. **Figure 8** Finger Knuckle from the hand dorsal surface.

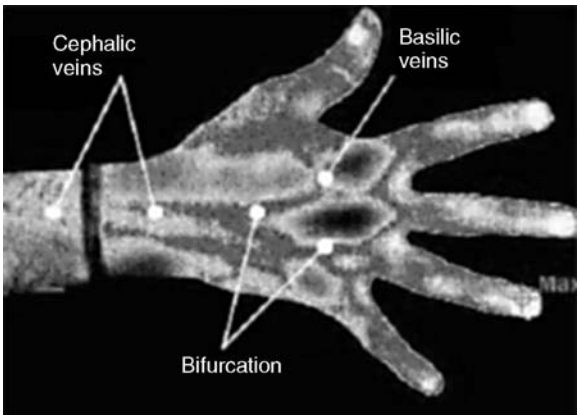
Hand Geometry

The anatomy of hand shape depends upon geometry of hand, length, width of fingers, and the span of the hand in different dimensions. The hand geometry biometric is not considered suitable for personal identification for the large scale user population as the hand geometry features are not highly distinctive. The requirement of the low cost imaging and low-complexity in feature extraction makes this biometric highly suitable for small scale applications (office attendence, building access etc.). The typical imaging setup for the acquisition of hand geometry images employ pegs to constrain the movement of fingers. However, recent publications have illustrated that the peg-free imaging can also be used to acquire images for hand geometry measurements. Such images can be used to extract length, width, perimeter, and area of palm/finger surface. These geometrical features of the hand can be simultaneously extracted with other biometric features, e.g. palmprint or fingerprint [10]. The anatomy of two human hands is quite similar and therefore hand geometry features from the left and right hands are expected to be similar. This is unlike the fingerprint or iris which shows characteristic distinctiveness in two separate (left and right) samples. The hand gestures also play very important anatomical representation in our daily life. Some of the examples of such activities are, waving the hand for familiar faces, making use of hands to call someone, representing the sign of victory with hands, fingers are used to point

someone, etc. The 3D hand gestures are a potential modality for gesture recognition and pose estimation and highly depend on the anatomy of individual's hands [11].

Hand Veins

Veins are hidden underneath the skin, and are generally invisible to the naked eye and other visual inspection systems. The pattern of blood veins is unique to every individual, even among identical twins. The veins are the internal structure responsible for carrying blood from one body part of the body to the other. Veins that are present in the fingers, palmar, and back of the hand surface are of particular interest in biometric identification. There are mainly two types of veins found on the dorsum of the hand, namely cephalic and basilic. Basilic veins are the group of veins attached with the surface of the hand. It generally consists of upper part of the back of hand. Cephalic veins are the group of veins attached with the wrist of the hand. It is often visible form the skin. The vein pattern of human hand can also be represented in the same way as fingerprint and palmprint by ridges and bifurcation points [12]. Figure 9 shows the vein structure on the back of human hand or on the palm dorsum surface. The spatial arrangement of the vascular network in the human body is stable and unique in individuals [13]. The prime function of vascular system is to provide oxygen to body parts. As the human body increases with age it extends or shrinks with the respective change in the body. Thus, the shape of hand vein changes with the physiological growth. During the adult life generally no major growth takes place and hence vein patterns are quite stable at the age of 20–50 years, at a later age the vascular system begins to shrink with the decline in the strength of bones and the muscles. These changes in vascular system make the vein pattern loose the earlier pattern. As the vascular system is a large and essential system of the body, it is largely affected by any change in the body; either by nature or by disease. Diabetes, hypertension, atherosclerosis, metabolic diseases, or tumors [14] are some diseases which affect the vascular systems and make it thick or thin.



Anatomy of Hand. Figure 9 Veins in the human hand.

The temperature of veins is quite different from its surrounding skin due to temperature gradient of skin tissues containing veins. This change in temperature can easily be observed in an image, captured by infra red thermal camera. However, such imaging is largely influenced by room temperature and humidity due to sensitivity of thermal cameras to these factors. Incorrect information about any of such factors can result in wrong approximation of temperature and affect the visibility of vein patterns. Based on the fact that the superficial veins have higher temperature than the surrounding tissue, the vein pattern at the back of the hand can be captured using a thermal camera. Other important aspect of vein anatomy relates to their spectral properties. Vein absorbs more infrared light as compared to its surrounding skin. This is due to level of blood oxygen saturation in the vein patterns. Therefore the vein pattern of a human hand can also be acquired using low-cost near infrared imaging [12]. The absorption and scattering property of infrared light depends upon exact wavelength used at the time of imaging, while at some wavelength arteries absorb more light than veins. Thus the same image of veins and arteries, acquired at low wavelength generates different intensity images.

The Reflectance Spectrum of Hand Skin

Besides the internal inherent structures that constitute the hand anatomy (as discussed above), some other biological properties of human hand can also be utilized to acquire unique features for biometric identification. One of such properties is the existence of distinguishing patterns in skin reflectance. The biological composition of skin and its response varies from individual to

individual. The spectral behavior of skin can be quantitatively measured as the ratio of light reflected over the incident light for a particular wavelength. This spectral analysis is one of the most reliable approaches to detect spoof biometric samples; as research in this area shows that the spectrum in the case of a mannequin is quite different from human skin [15]. It is important to note that the spectral characteristics of the palm are quite identical to that of back of hand with little increase in wavelength due its reddish color. The spectral reflectance of skin is quite independent of any particular race or species and therefore it cannot be used for any such classification. However, the darker skin reflects smaller proportion of incident light, therefore the variation in curvature is also low [15].

Summary

The structure of human hand is quite complicated and consists of a variety of soft tissues and bones. The hand based biometrics system exploits several internal and external features that are quite distinct to an individual. However, some features or traits have been observed to be highly stable while some are more conveniently acquired (e.g. hand geometry). The individuality in the uniqueness of the hand based bio metrics is highly dependent on the intrinsic anatomical properties of the hand. There has been very little work to explore several anatomical characteristics of hand, e.g. muscles, nerves, etc., for biometric identification. The success of a biometric modality highly depends on its uniqueness or the individuality, which can be better explored from the human anatomy and the biological process that generates corresponding physiological characteristics.

Related Entries

- ▶ Palmprint, 3D
- ▶ Hand Geometry
- ▶ Hand Vein
- ▶ Palmprint Features

References

1. Pedro Beredjiklian, M.D.: The Structure of the Hand. http://files.ali-aba.org/thumbs/datastorage/skoob/articles/BK40-CH11_thumb.pdf (2003)
2. Kumar, A., Wong, D.C., Shen, H.C., Jain, A.K.: Personal verification using palmprint and hand geometry biometric. In: Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 668–675. Proceedings of Fourth International Conference, Guildford, UK (June 9–11, 2003)
3. Norman, W.: The Anatomy Lesson. http://home.comcast.net/~WNOR/lesson5mus&tendonsofhand.htm (1999)
4. Richards, L., Loudon, J.: Hand Kinesiology. The University of Kansas, http://classes.kumc.edu/sah/resources/handkines/nerves/dermatome.htm (1997)
5. Kanhangad, V., Zhang, D., Nan, L.: A multimodal biometric authentication system based on 2D and 3D palmprint features. In: Biometric Technology for Human Identification, vol. 6944, pp. 69,440C–69, 440C. Proceedings of SPIE, Orlando, Florida (March 2008)
6. Jain, A.K., Demirkus, M.: On Latent Palmprint Matching. MSU Technical Report (2008)
7. Ashbaugh, D.R.: Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology. CRC, Boca Raton (1999)
8. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003)
9. Jain, A.K., Prabhakar, S., Pankanti, S.: On the similarity of identical twin fingerprints. Pattern Recogn. **35**, 2653–2663 (2002)
10. Kumar, A., Ravikanth, Ch.: Personal authentication using finger knuckle surface, IEEE Trans Information Forensics & Security, **4**(1), 98–110 (2009)
11. Guan, H., Rogerio, F.S., Turk, M.: The isometric self-organizing map for 3D hand pose estimation. In: Automatic Face and Gesture Recognition, pp. 263–268. The Seventh International Conference, Southampton, UK (2006)
12. Kumar, A., Prathyusha, K.V.: Personal authentication using hand vein triangulation, vol. 6944, pp. 69, 440E–69, 440E–13. Proceedings of SPIE Conference Biometric Technology for Human Identification, Orlando, Florida (Mar. 2008)
13. Nadort, A.: The Hand Vein Pattern Used as a Biometric Feature. Master's thesis, Netherlands Forensic Institute, Amsterdam (2007)
14. Carmeliet, P., Jain, R.K.: Angiogenesis in cancer and other diseases. Nature **407**, 249–257 (2000)
15. Angelopoulou, E.: The Reflectance Spectrum of Human Skin. Tech. Rep. MS-CIS-99-29, University of Pennsylvania (1999)

Analysis-by-Synthesis

Analysis by synthesis is the process that aims to analyze a signal or image by reproducing it using a model. The objective is to find the value of the model parameters that synthesize the closest image possible in the span of the model. It is then an optimization problem that

requires the setting of a cost function (e.g., sum of squares) and of a model with a small number of parameters. The model must be able to generate typical variations (such as pose, illumination, identity and expression for face images), to enable the analysis of a signal or image that includes expected variations.

The analysis-by-synthesis approach of heterogeneous face matching compares between an enrollment image A and an image A_0 which is synthesized from an input probe image in such a way that the image properties of A_0 resemble those of A.

► Face Sample Synthesis
► Heterogeneous Face Biometrics

Analytic Study

An analytic study is one where the goal is the utilization of the information gathered for improvement of the process going forward as opposed to an enumerative study.

► Test Sample and Size

And-Or Graph

An And-Or graph is a 6-tuple for representing an image grammar G

$$G_{\text{and-or}} = S, V_{\text{N}}, V_{\text{T}}, R, \Sigma, P$$

where S is a root node for a scene or object category, V_{N} is a set of non-terminal nodes including an And-node set and an Or-node set, V_{T} is a set of terminal nodes for primitives, parts and objects, R is a number of relations between the nodes, Σ is the set of all valid configurations derivable from the grammar, and P is the probability model defined on the And-Or graph.

► And-Or Graph Model for Faces

And–Or Graph Model for Faces

Feng Min[1,2], Jinli Suo[2,4], Song-Chun Zhu[2,3]
[1]Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, China
[2]Lotus Hill Institute for Computer Vision and Information Science, China
[3]Departments of Statistics and Computer Science, University of California, Los Angeles
[4]Graduate University of Chinese Acadamy of Sciences, China

Synonym

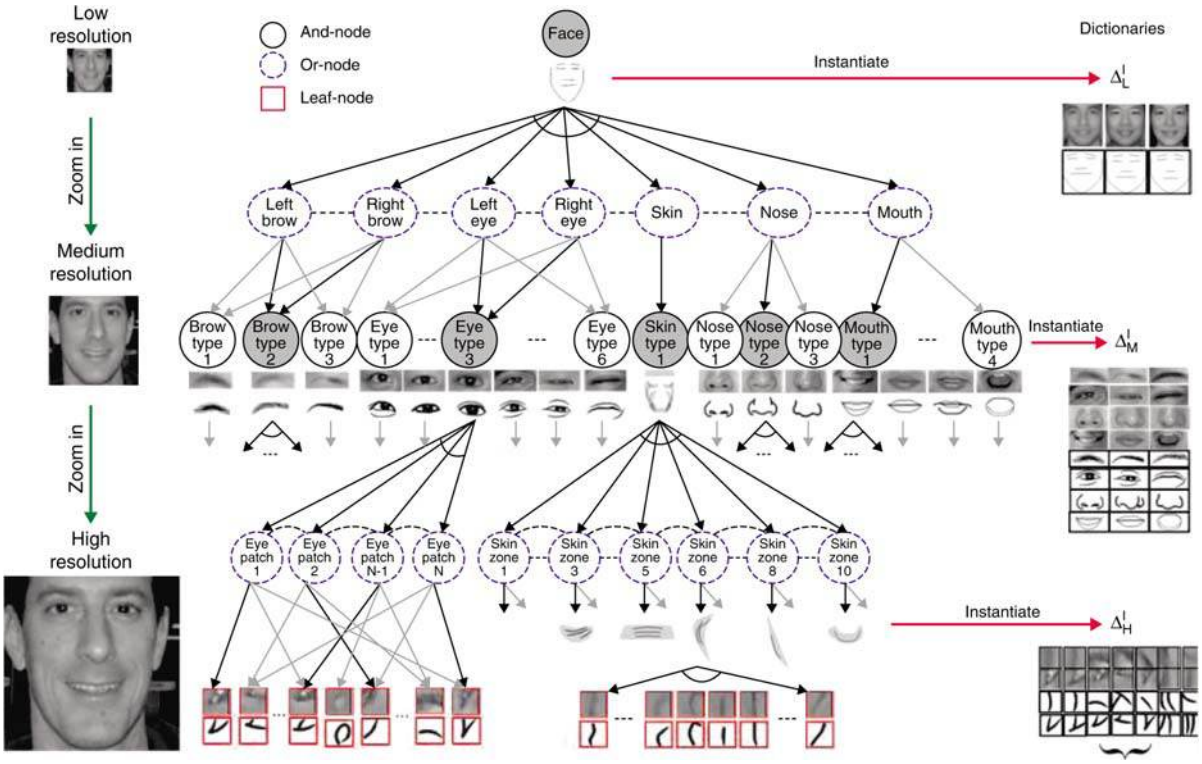
And–Or Graph Model

Definition

For face modeling, an ► And–Or graph model was first proposed in [1] as a compositional representation for high resolution face images. In an And–Or graph, the And nodes represent coarse-to-fine decompositions and the Or-nodes represent alternative components for diversity. The And–Or graph face model, as illustrated in Fig. 1, has three levels: the first level describes the general appearance of global face and hair; the second level refines the representation of the facial components (eyes, eye brows, nose, mouth) by modeling the variations of their shapes and subtle appearance; and the third level provides further details of the face components and divides the face skin into nine zones where the wrinkles and speckles are represented. The And–Or graph provides an expressive model for face diversity and details, and thus is found to be especially efficient for applications in ► face sketching generation and ► face aging simulation.

Introduction

Human faces have been extensively studied in computer vision and graphics for their wide applications: detection, recognition, tracking, expression recognition, and nonphotorealistic rendering (NPR). Many face models have been proposed, for example,



And–Or Graph Model for Faces. Figure 1 An illustration of the compositional And–Or graph representation of human face. The left column is a face image at three resolutions. All face images are collectively modeled by a three-level And–Or graph in the middle column. The And nodes represent decomposition and the Or nodes represent alternatives. Spatial relations and constraints are represented by the horizontal links between nodes at the same level. By the selection of alternatives, the And–Or graph turns into a *parse graph* for a face instance. The right column represents the dictionaries at three scales: Δ_H^I, Δ_M^I, *and* Δ_L^I. *From Xu* et al. [1].

EigenFace [2], FisherFace [3], Laplacianfaces [4] and their variants [5], deformable templates [6], the active shape models, and active appearance models [7, 8, 9]. Most of these models are mainly used for face detection, localization, tracking, and recognition.

Although these face models have achieved reasonable successes in face detection, recognition, and tracking, they use templates of fixed dimensions at certain low-middle resolutions, and thus are limited by their expressive powers in describing facial details in higher resolutions, for example, subtle details in the different types of eyes, nose, mouths, eyebrows, eyelids, muscle relaxations due to aging, skin marks, motes, and speckles. Consequently, these models are less applicable to applications that entail high precision, such as face sketch generation and face aging simulation. For the latter tasks, Xu et al. [1] proposed a compositional And–Or graph representation for high-resolution face images. Adopting a coarse-to-fine hierarchy with the Or nodes representing the alternatives, the And–Or graph can represent a large diversity of human faces at different resolutions.

Compositional And–Or Graph Representation for Faces

A compositional And–Or graph describes all types of faces collectively at low, medium, and high resolutions, as shown in Fig. 1. There are three types of nodes in the And–Or graph: And-nodes, Or-nodes and leaf-nodes. An And-node either represents a way for decomposition at higher resolution or terminates in a Leaf-node at lower resolution. An Or-node stands for a switch pointing to a number of alternatives components. For example, an Or-node of eye could point to different types of eyes. A leaf-node is an image patch or

image primitive with shape and appearance attributes. The And–Or graph has horizontal lines (see dashed) to specify the spatial relations and constraints among the nodes at the same level. By choosing the alternatives at Or nodes, the And–Or graph turns into an And-graph representing a face instance, which is called a *parse graph*. Thus, the And–Or graph is like a "mother template," which produces a set of valid face configurations, each of which is a composition of the image patches or primitives at its leaf nodes.

At low resolutions, the face is represented as a traditional Active Appearance Model (AAM) [8], which describe the general face shape, skin color, etc. At medium resolutions, the face node expands to a number of Or-nodes for facial components (eyebrows, eyes, nose, and mouth) and skin zone. For each component, a number of AAM models are used as the alternatives for the Or node. At high resolutions, the nodes of facial component and skin zone further expand into a number of Or-nodes describing the local structure of components and free curves (wrinkles, marks, etc.) in detail.

Model Computation

For an input high-resolution face image, the algorithm computes a parse graph in a Bayesian framework in three levels from coarse to fine.

At the first level, the face image is down-sampled, and the algorithm computes the AAM-like representation W_L with global transform T, geometrical deformation α_{geo}, and photometric appearance β_{pht} by maximizing the posterior probability,

$$\begin{aligned} W_L &= \arg\max\ p(I_L^{\text{obs}}|W_L;\Delta_I)p(W_L) \\ W_L &= (T,\alpha_{\text{geo}},\beta_{\text{pht}}). \end{aligned} \tag{1}$$

At the second level, a number of AAM-like models are trained for each facial component. The algorithm takes a down-sampled medium resolution face image and W_L as the input and conducts a constrained search for W_M conditioned on W_L. The variables are computed by maximizing the posterior probability,

$$\begin{aligned} W_M &= \arg\max\ p(I_M^{\text{obs}}|W_L,W_M;\Delta_I,\Delta_I^{\text{CP}}) \\ &\quad p(W_M|W_L) \\ W_M &= (l^i,\alpha^i_{l^i,\text{geo}},\beta^i_{l^i,\text{pht}})_{i=1}^{6} \end{aligned} \tag{2}$$

At the third level, the face area is decomposed into zones that refine the sketches of local structures, based on the searching results at medium resolution level. The variables at this layer are inferred by maximizing the posterior,

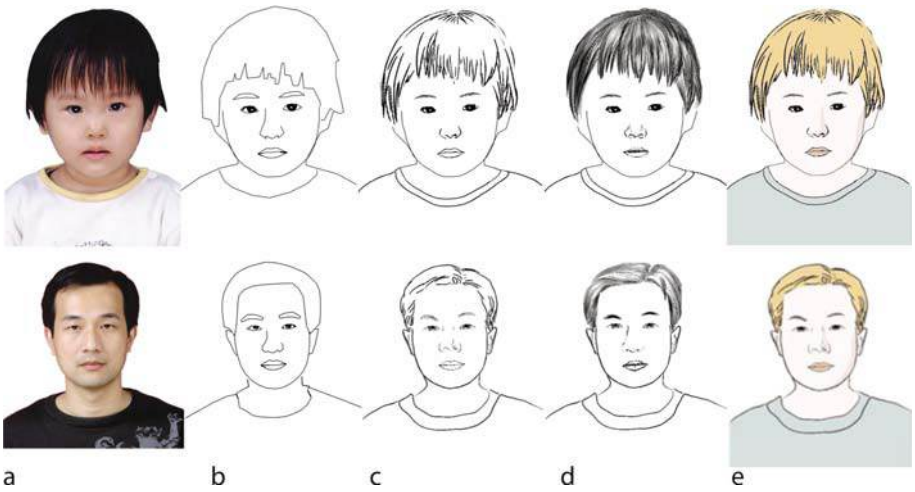
$$\begin{aligned} W_M &= \arg\max\ p(I_H^{\text{obs}}|W_M,W_H;\Delta_I^{\text{CP}},\Delta_I^{\text{SK}}) \\ &\quad p(W_H|W_M) \\ W_H &= (K,\{(l_k,t_k,\alpha_k):k=1,2,...,K\}) \end{aligned} \tag{3}$$

Applications

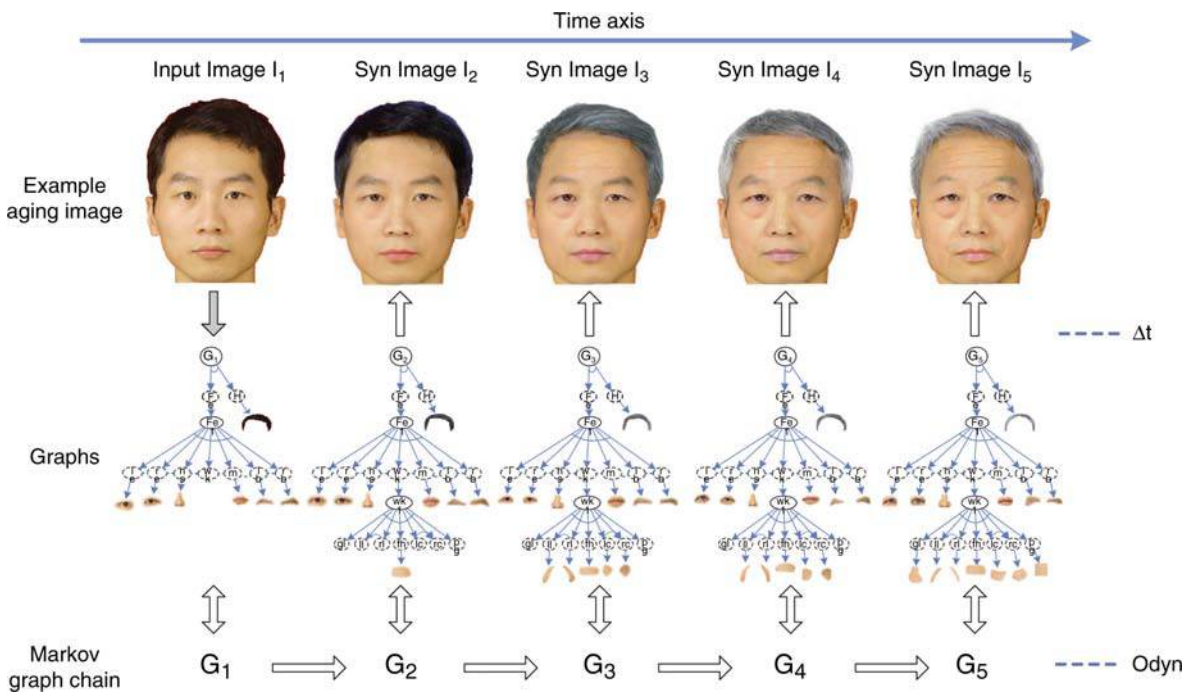
The And–Or graph face model has been applied to two applications: automatic face sketch and portraiture generation in [11] and face aging simulation in [10].

Min et al. [11] developed an automatic human portrait system based on the And–Or graph representation. The system can automatically generate a set of life-like portraits in different artistic styles from a frontal face image as shown in Fig. 2. The And–Or graph is adopted to account for the variabilities of portraits, including variations in the structures, curves, and drawing style. Given a frontal face image, a local AAM search is performed for each facial component, based on the search result, the hair and collar contours can be inferred. Then, using predefined distances, a template matching step finds the best matching template from sketch dictionaries for each portrait component. Finally, the strokes of specific style will render each component into stylistic results. Making good use of the large sketch dictionaries in different styles, it can conveniently generate realistic portraits with detailed face feature of different styles.

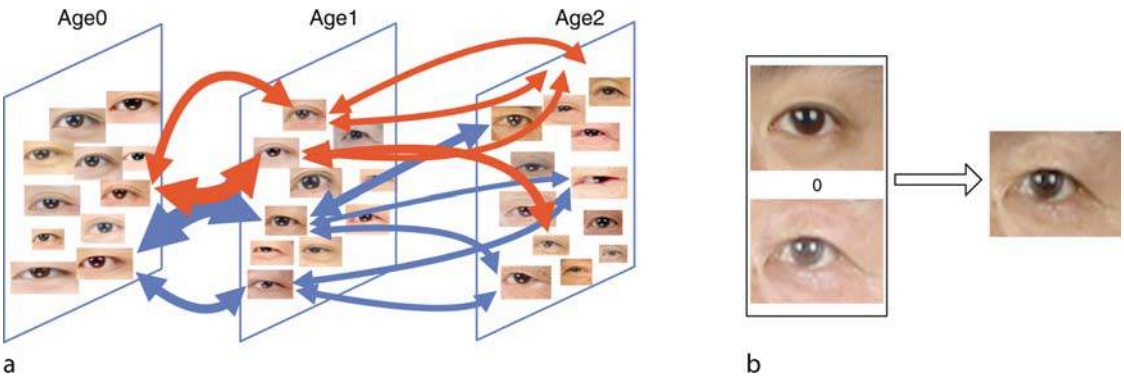
Suo et al. [10] augmented the compositional face model [1] with aging and hair features. This augmented model integrates three most prominent aspects related to aging changes: global appearance changes in hair style and shape, deformations and aging effects of facial components, and wrinkles appearance at various facial zones. Then face aging is modeled as a dynamic Markov process on this graph representation, which is learned from a large dataset. Given an input image, the aging approach first computes the parse graph representation, and then samples the graph structures over various age groups according to the learned dynamic model. Finally the sampled graphs generate face images together with the dictionaries. Figure 3 is an illustration of the dynamic model for aging over the parse graphs. I_1 is an input



And–Or Graph Model for Faces. **Figure 2** The result of applying compositional And–Or graph model to portraiture generation. (**a**) is an input frontal face image, (**b**) is a draft sketch obtained by image processing methods based on AAM search result of face contour and face component, (**c**)–(**e**) are separately three rendered results by the sketch dictionaries in literary, pencil, and colored style. *From Min* et al. [11].



And–Or Graph Model for Faces. **Figure 3** An aging process can be modeled by a Markov Chain on the parse graphs G_t where t is an age period. The first row is an aging sequence of face, I_1 is the input image, and the other four are simulated aged images. The second row is the graph representations of the image sequence. Third row is the corresponding parse graphs G_t, which form a Markov Chain. $\Theta_{img,t}$ includes the parameters for generating the images from G_t and Θ_{dyn} the parameters for aging progression. *From Suo* et al. [10].



Annotated Face Model. Figure 4 (**a**) is a training subset for dynamic learning of face aging. (**b**) is one simulated result of eye aging. *From Suo* et al. [10].

young face image and G_1 is its graph representation. I_2 to I_5 are four synthesized aged images in four consecutive age groups generated by G_2 to G_5. $\{G_1, G_2, \ldots, G_5\}$ is a chain of parse graphs describing face aging procedure.

The compositional model decomposes face into parts, and this strategy provides the potential of learning the statistics of each node separately. In Suo et al. [10], aging dynamics are learned from similar parts cropped from different persons, Fig. 4(a) gives a training subset of eye aging and (b) is the aging results. Human experiments have validated that this aging process is perceptually plausible.

Summary

The compositional And–Or graph model is an expressive representation of high-resolution human face. With the selection of alternatives at Or nodes, the And–Or graph can model the large diversity of different faces as well as the artistic styles. The decomposition allows learning of parts and the spatial constraints, and alleviates the difficulty of training set collection. The model has been applied to automatic portrait generation and face aging simulation. The authors' argue that the model should also improve other applications such as face recognition and expression analysis.

Related Entries

- ▶ And–Or Graph
- ▶ Face Aging
- ▶ Face Sketching

References

1. Xu, Z.J., Chen, H., Zhu, S.C.: A high resolution grammatical model for face representation and sketching. CVPR (2005)
2. Turk, M., Pentland, A.: Eigenfaces for recognition. CVPR (1991)
3. Belhumeur, P., Hespanda, J., Kriegman, D.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. PAMI **19**(7), 711–720 (1997)
4. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.: Face recognition using Laplacianfaces. PAMI **27**(3), 328–340 (2005)
5. Kim, J., Choi, J., Yi, J., Turk, M.: Effective representation using ICA for face recognition robust to local distortion and partial occlusion. PAMI **27**(12), 1977–1981 (2005)
6. Yuille, A.L., Cohen, D., Hallinan, P.: Feature extraction from faces using deformable templates. IJCV **8**(2), 99–111 (1992)
7. Cootes, T.F., Taylor, C.J., Cooper, D., Graham, J.: Active shape models-their training and application. Comput. Vison Image Understand **61**(1), 38–59 (1995)
8. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. Proc. of ECCV (1998)
9. Cootes, T.F., Taylor, C.J.: Constrained active appearance models. Proc. of ICCV (2001)
10. Suo, J.L., Min, F., Zhu, S.C.: A multi-resolution dynamic model for face aging simulation. CVPR (2007)
11. Min, F., Suo, J.L., Zhu, S.C., Sang, N.: An automatic portrait system based on And–Or graph representation. EMMCVPR (2007)

Annotated Face Model

The annotated face model (AFM) is a 3D model of a human face. The AFM defines the control points of subdivision surfaces and it is annotated into different

areas (e.g., mouth, nose, eyes). Using a global parameterization of the AFM, the polygonal representation of the model can be converted into an equivalent geometry image representation.

► Face Recognition, 3D-Based

Anthropometry

Anthropometry is the study of human body measurements for use in anthropological classification and comparison. It has been used to assess nutritional status, to monitor the growth of children, and to assist in the design of office furniture and garment.

► Background Checks

Anthroposcopy

Anthroposcopy is about visual observation of the human body such as skin color, body shape, in contrast to more objective and precise anthropometry which is about the measurement of the hyman body.

► Gait, Forensic Evidence of

Anti-Spoofing

A biometric spoof is an artificials mimic of a real biometric. Anti-spoofing is a technical measure against biometric spoofing. Liveness detection is one of such techniques.

► Biometric Liveness
► Biometric Spoofing Prevention
► Liveness Detection: Fingerprint
► Liveness Detection: Iris

Appearance-Based Gait Analysis

Gait analysis by using information contained in an image, with or without using temporal information.

► Gait Recognition, Motion Analysis for

Application Programming Interface (API)

An API is a set of software functions by which a software application can make requests of a lower level software service, library, or Operating System (OS). It is a way for one piece of software to ask another piece of software to do something. In the case of OS calls, the application may request basic functions such as file system access. Other APIs are more specific to the servicing software. For example, BioAPI is a set of biometric programming functions that can be used to develop a biometric system.

► Interfaces, Biometric
► Large Scale System Design

Artifact

An artifact is a man-made object or device; in connection to biometrics, artifacts are man-made imitations of biometric traits to circumvent a biometric system. An example of an iris artifact is a contact lens with printed or hand-painted iris patterns.

► Liveness Detection: Iris

Artificial Biometrics

► Biometric Sample Synthesis

Artificial Digital Biometrics

▶ Biometric Sample Synthesis

Artificial Fingerprints

▶ Fingerprint Sample Synthesis

Artificial Image Biometrics

▶ Biometric Sample Synthesis

ASN.1

▶ Abstract Syntax Notation one

Asset Protection

▶ Transportable Asset Protection

Association

▶ Human Detection and Tracking

Attack Trees

An attack tree is a diagram which graphically shows the conceptual structure of a threat on a computer system. It was designed by Bruce Schneier [(1999) Attack Trees. Dr. Dobb's Journal] to help organize analysis of system security. Attack trees are multi-level diagrams with one root and leaves, and children. Each node describes a condition which is either necessary or sufficient to enable the node above. For example, the attack (root node) "Open Safe", may occur due to "Pick Lock," "Learn Combo", or "Cut Open Safe". The node "Learn Combo" may, in turn, occur due to nodes "Eavesdrop" or "Bribe", which in turn depend on further factors. Further analysis of the attack tree may be performed by assigning each block a parameter (feasibility, required technical skill, expense) and calculating the cost for the overall attack.

▶ Biometric Vulnerabilities, Overview

Audio–Visual-Dynamic Speaker Recognition

▶ Lip Movement Recognition

Audio-Visual Fusion

Audio-visual fusion combines audio and visual information to achieve higher person recognition performance than both audio-only and visual-only person recognition systems. There exist various fusion approaches, including adaptive approaches, which weight the contribution of audio and video information based on their discrimination ability and reliability.

▶ Lip Movement Recognition

Audio-Visual Speaker Recognition

In audio-visual speaker recognition, speech is used together with static video frames of the face or certain parts of the face (face recognition) and/or video sequences of the face or mouth area to improve person recognition performance. The main advantage of audio-visual

biometric systems lies in their improved robustness, and resilience to spoofing. Each modality can provide independent and complementary information and therefore, prevent performance degradation due to the noise present in one or both of the modalities.

► Lip Movement Recognition

Audio-Visual Speech Processing

Under some circumstances, such as in very noisy environments, it could be useful to use not only the acoustic evidence of the speech, but also visual evidence by recording the movement of the lips and processing both evidences together. This processing of audio and visual speech is commonly referred to as audio-visual speech processing.

► Voice Device

Authentication

Biometric authentication is a synonym for biometric recognition, meaning either verification or identification in biometrics.

► Biometrics, Overview
► Fraud Reduction, Overview
► Verification/Identification/Authentication/Recognition: The Terminology

Authentics Distribution

The probability distribution of the match score a biometric for cases where one instance of a biometric template is compared against another instance derived from the same individual as the first.

► Iris on the Move

Automated Fingerprint Identification System

A computerized system that acquires, stores, and manages a large scale fingerprint database and criminal history, and provides fingerprint search service for biometric identification and fraud prevention.

► Fingerprint, Forensic Evidence of
► Fingerprint Matching, Automatic

Automatic Classification of Left/Right Iris Image

Yung-hui Li[1], Marios Savvides[2]
[1]Language Technology Institute, Carnegie Mellon University, Pittsburgh, PA, USA
[2]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Synonym

Mislabeled Iris data correction

Definition

Many iris acquisition devices capture irises from a single eye at a time. The device operator must typically enter meta-data such as name, address, and which eye by hand. In many deployment scenarios it is easy for the device operator to be distracted and mislabel the eyes. Such mislabeling can pose serious problems for database indexing. In this article, the authors describe an extremely efficient algorithm for automatic classification of eyes into left/right categories. This algorithm makes use of the iris/pupil segmentation information that is already computed for most iris recognition algorithms, so it poses a minimal computational load and requires minimal modifications to existing iris recognition systems.

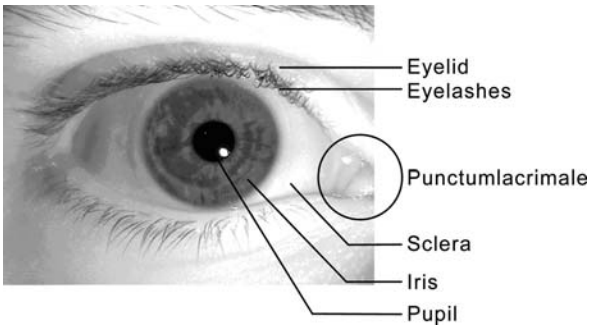
Introduction

Iris recognition is generally considered to be one of the most effective biometric modalities for biometric identification [1]. Iris is a good biometric because (1) the iris is rich in texture and that texture has many degrees of freedom [2]; (2) the iris is both protected and accessible; (3) the iris texture is thought to be stable throughout most of a person's life, barring catastrophic injury, or illness; (4) the fraction of a the population that cannot present an iris due to injury or congenital defect such as aniridia is small; (5) the iris can be easily accessed in a non-contact manner from moderate distances.

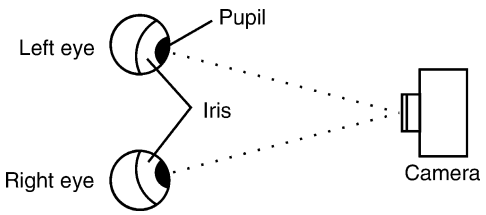
The performance and reliability of all biometric identification systems depend crucially on the quality of the enrollment data. Many real-world application scenarios use single-iris acquisition devices that are prone to mislabeling of left versus right iris due to human error. If the enrollment database has been corrupted by such errors, it is necessary to search both left and right eyes during verification or identification. For iris recognition algorithms in general, this results in roughly a factor of two increases in the computational cost of the search – for any individual for whom the enrollment images are swapped. Hence, scrubbing enrollment databases of such errors will be useful as long as the automatic procedure has an error rate which is smaller than the error rate of the human device operators. The authors have developed an algorithm that classifies iris images into left/right categories based on an analysis of the pupil and iris segmentation data that is already available in the pre-processing stage of most iris recognition algorithms. Hence, it does not introduce an increase in the computational load and can provide significant increase in the search efficiency for enrollment databases that have left/right classification errors.

Basis of the Algorithm

Figure 1 shows an example of a right eye image in which the main components of the eye can be seen: pupil, iris sclera, eyelid, and eyelashes. Note the ▶ punctum lacrimale, the D-shaped corner where the upper eyelid meets lower eyelid. For right eye images this is always on the right and for a left eye images, it is on the left. The location of the punctum lacrimale is one of the most effective ways for humans to distinguish left eyes from right eyes.



Automatic Classification of Left/Right Iris Image.
Figure 1 A right eye image.

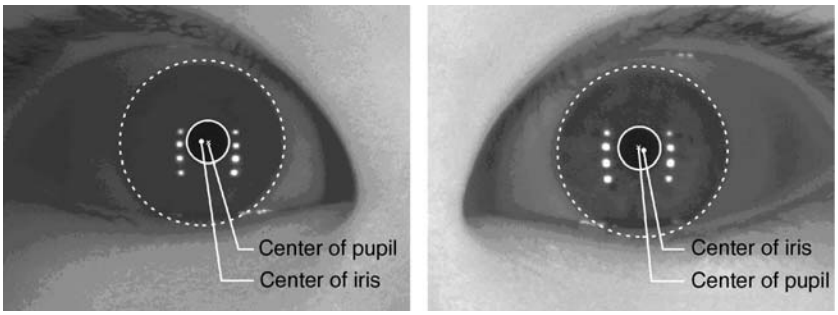


Automatic Classification of Left/Right Iris Image.
Figure 2 Illustration of the relative location of eye balls and camera.

Though the punctum lacrimale is easily distinguished by humans, it is currently a challenging feature for machine vision systems. Furthermore, the punctum lacrimale is not always visible in iris recognition images (due to sensor acquisition and processing) and in some cases it is not particularly prominent which causes difficulty even for humans. Other eye-shape characteristics have been used for left/right classification, but all of them suffer from the same sorts of inherent person to person variability of eye shape. To avoid these problems, the authors chose to consider analysis of simpler iris image characteristics: the geometric locations of the pupil and iris that are already computed by most iris recognition algorithms. This analysis can be understood with the help of Fig. 2, an illustration of the relative positions of the left eye, right eye and a camera at a particular location; the dotted line denotes the line of sight for each of the two eyes.

It is observed that the distance between the camera and the eyes is not infinity; the two lines of sight are not parallel to each other. Rather, they intersect with each other at an angle.

Figure 3 shows images of left and right eyes looking at the same camera. It is seen that the pupil in the left



Automatic Classification of Left/Right Iris Image. Figure 3 (**a**) An example image of right eye with center of pupil and iris marked with x and ., respectively. (**b**) An example image of left eye with center of pupil and iris marked with x and respectively.

eye is located closer to its punctum lacrimale (i.e., to the right of the iris center) and the pupil of the right iris is similarly located closer its punctum lacrimale (i.e., to the left of the iris center). These observations (backed with empirical evaluations) can provide a simple, elegant, yet very effective left/right classifier for iris images. Note that the relative position of the center of the pupil and iris: if the center of the pupil is on the left side of the center of the iris, it is more likely that this is a left iris eye; otherwise, it is more likely to be a right iris eye.

The algorithm can be summarized as:

1. Perform ▶ iris localization. This can be done by using any kind of iris segmentation algorithm commercially or available academically, as long as it is effective, precise [2–11] and provides the (x, y) image coordinates of the pupil and iris.
2. Retrieve the x-coordinate of both the center of pupil and iris.
3. If the x-coordinate of the center of pupil is smaller than x-coordinate of the center of iris, classify it as left eye image.
4. If the x-coordinate of the center of pupil is larger than x-coordinate of the center of iris, classify it as right eye image.
5. If the x-coordinate of the center of pupil is exactly the same as x-coordinate of the center of iris, then a decision can not be made with this algorithm and input from other pattern classifiers (e.g., punctum lacrimale detector, eye-shape analysis, or random guess) can be used.

Since this method is used for the existing localization/segmentation data that is extracted in most deployed iris recognition systems it poses a minimal additional computational load. The precision of the algorithm vastly depends on the accuracy of the iris localization process. In summary, the proposed methodology does not add any overhead to existing iris processing framework, and can be executed extremely fast and can be used as a meta-analysis tool to rectify already acquired datasets. In the case that a decision cannot be made, the tool can prompt the user to manually see if he can identify the eye or further feature detectors such as punctum lacrimale detectors or eye-shape analysis can be used to make the decision and in some cases, a random assignment might be acceptable.

Algorithm Performance

The authors evaluated the algorithm on NIST's Iris Challenge Evaluation (ICE) 2005 database [12]. The ICE 2005 dataset contains a total of 2953 irises made up from 1528 left irises and 1425 right irises captured from an LG EOU 2200 single iris capture unit. For these experiments, the images were segmented using the segmentation algorithm described in. One of the 2953 irises was incorrectly segmented; that image had an iris that was badly off-axis nature – an outlier in this dataset; that image was omitted from the subsequent analysis.

The segmentation data was analyzed using the left/right classification algorithm; the results are shown in Table 1. The data for all images and the left and right eyes is presented separately. The error rates are below 1%. The fraction of images for which a decision could not be made is ~6%.

Automatic Classification of Left/Right Iris Image. **Table 1** The experimental results for Left vs. Right eye classification, on ICE 2005 database

Category	Image count	Misclassified	Undecided	Correct identification rate
All images	2953	27	172	99.1%
Left	1528	15	78	99%
Right	1425	12	94	99.2%

If the algorithm is used for purely automated classification and assigned the undecided cases at random, a classification error rate of the order of approximately 4% would be achieved. Only 5.8% of the images were not able to be classified as left or right, and the ability to determine this is crucial as it allows either to prompt human input or to employ further more complex feature extraction to see if a determination can be done. However from the 94% of the images that were automatically determined that a classification decision could be made to whether they belong to left or right irises, the authors' classification algorithm made the correct label assignment 99.1% of the time which is a significant achievement of the proposed algorithm. The implications of this approach is that it allows to reduce the computational search time of matching by a factor of 2 by applying a fully automatic method to partition left/right iris datasets. For an iris recognition system with a large database with high loading, this could result in a substantial reduction in the cost of the server farm needed to support the system.

Summary

Automatic left/right classification of iris images is important in iris recognition systems. The authors have presented a simple algorithm for automatic classification that is efficient, effective and introduces minimal additional computational load on the system. Experimental test on the ICE 2005 database demonstrate that the algorithm can provide fully automated classification and has the ability to determine when it is not confident to make a correct classification decision, on the ICE dataset this was approximately 5.8% of the data where it determined that further human input or other feature extraction processing is necessary. On the remaining 94.2% of the images that it determined a decision could be made, it achieved a correct classification rate of 99.1% on labeling the images as left or right irises. This can provide a roughly 2× reduction in the computational load for irises matching in large databases.

Related Entries

- ▶ Image Pattern Recognition
- ▶ Iris Databases
- ▶ Iris Recognition, Overview

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circ. Syst. Video Tech. 14(1) (2004)
2. Daugman, J.: How iris recognition works. IEEE Trans. Circ. Syst. video Tech. 14(1) (2004)
3. Wildes, R.: Iris recognition: An emerging biometric technology. Proceedings of IEEE **85**, 1348–1363 (1997)
4. Huang, Y.P., Luo, S.W., Chen, E.Y.: An efficient iris recognition system. International Conference of Machine Learning and Cybernetics **1**, 450–454 (2002)
5. Liu, Y., Yuan, S., Zhu, X., Cui, Q.: A practical iris acquisition system and a fast edges locating algorithm in iris recognition. In: IEEE Instrumentation and Measurement Technology Conference, Vol 26, pp. 166–168 (2003)
6. Sung, H., Lim, J., Park, J.H., Lee, Y.: Iris recognition using collarette boundary localization. International Conference on Pattern Recognition **IV:** 857–860 (2004)
7. Masek, L.: Recognition of human iris patterns for biometric identification. Master thesis, University of Western Australia, 2003. Available from: http://www.csse.uwa.edu.au/verb++pk/studentprojects/libor/
8. Liu, X., Bowyer, K.W., Flynn, P.J.: Experiments with an improved iris segmentation algorithm. Fourth IEEE Workshop on Automatic Identification Technologies, Vol 113, 118–123 (2005)
9. Pan, L., Xie, M.: The algorithm of iris image processing. Fourth IEEE Workshop on Automatic Identification Technologies 134–138 (2005)
10. He, X.F., Shi, P.F.: A novel iris segmentation method for hand-held capture device. Springer LNCS 3832: International Conference on Biometrics 479–485 (2006)
11. Feng, X., Fang, C., Ding, Z., Wu, Y.: Iris localization with dual coarse-to-fine strategy. International Conference on Pattern Recognition 553–556 (2006)

12. National Institute of Standards and Technology. Iris challenge evaluation 2005 workshop presentations. Available from: http://iris.nist.gov/ice/presentations.htm. Last Accessed March 24, 2009
13. Matey, J.R., Naroditsky, O., Hanna, K., Kolczynski, R., LoIacono, D., Mangru, S., Tinker, M., Zappia, T., Zhao, W.Y.: Iris on the Move: acquisition of images for iris recognition in less constrained environments. In: Proceedings of IEEE **94**(11), 1936–1946 (2006)
14. Thornton, J., Savvides, M., Vijaya Kumar, B.V.K.: A Bayesian Approach to Deformed Pattern Matching of Iris Images. IEEE Trans Pattern Anal Mach Intell **29**(4), 596–556 (2007)

Average Correlation Energy (ACE)

The result of applying a correlation filter is a two-dimensional correlation plane. The average energy of this plane can be estimated by first summing up the squares of the values on every pixel and then dividing it by the total number of the pixels. This value is called average correlation energy. It is important to minimize this term in the design of correlation filters because it represents the average height of the sidelobes on the correlation plane. If we would like to see a sharp peak on the correlation plane for the authentic comparison, criteria for minimizing sidelobes has to be added into the optimization process.

► Iris Recognition Using Correlation Filters

Azimuth

In a plane, the angle measured clockwise from a coordinate axis and a line, with values ranging from 0° to 360°. This measure is a component of the pen orientation in handwriting capture devices.

► Signature Features

B

Background Checks

Peter T. Higgins
Higgins, Hermansen, Banikas, LLC, VA, USA

Synonyms

Vetting; Credit check; Personal information search; Preemployment screening; Disclosure check; Criminal record search; Criminal history check

Definitions

There are multiple types of Background checks that all involve reviewing past, recorded behavior:

1. Job Applicants: When people apply for a position of trust (e.g., a school teacher, a lawyer, or bank teller) a background check is part of the way of determining if the applicant is suitable – a positive result. These are known as Applicant or Civil Background Checks. In the US:
   (a) If there is a state or federal law requiring a national check then applicant fingerprints (with minimal supporting biographic information) can be submitted to the Federal Bureau of Investigation (FBI) for a search.
   (b) If the applicants are seeking federal employment then their biographic data and fingerprints can be submitted to the FBI for a search.
   (c) If the applicants are applying for a job not covered by a state or federal law they are restricted to commercial background checking services – companies that have aggregated financial, court, motor vehicle and other records.
2. Applicants for Credit: When people apply for credit cards or a large financial commitment (e.g., a mortgage) a background check is part of the way of determining if the applicant is suitable – a positive result. These are known as Credit Checks and are performed by commercial background checking services.
3. Applicants for Government Benefits: When people apply for a visa, passport, drivers license, social security and other benefits governments use varying levels of background checks to weed out fraud (e.g., multiple applications with different identities but for the same subject), previously denied persons, etc. Other than possibly checks for visas, most of these checks say nothing about the suitability of a person for trustworthiness.
4. Criminals: In the criminal justice community Background Checks are used when a person is arrested to determine if an arrestee already has a criminal record that they are hiding – a negative result that will be used in setting bail, sentencing, etc. In the US:
5. Arrestee fingerprints (with minimal supporting biographic information) can be submitted to the FBI for a search.

Background Checks are primarily based on textual information (e.g., name and date of birth) searches of bank, court, credit card issuer, and other files or textual searches and in some cases are combined with biometric based (e.g., fingerprint) searches of criminal or undesirable persons (e.g., persons previously deported) records.

The ANSI/IAI 2-1988 American National Standard for Forensic Identification Glossary of Terms and Acronyms defines ▶ "criminal history check" as "*A search of name indices and/or fingerprint files to determine whether or not a subject has a prior criminal record.*"

The same American National Standard glossary defines "criminal history" as "*A chronological summary of an individual's criminal activity which may include the dates of the activity, the individual's name, aliases and other personal descriptors, the identities of the reporting agencies, the arrest charges, dispositions, etc.*"

© 2009 Springer Science+Business Media, LLC

The UK Criminal Records Bureau performs an Enhanced Disclosure Check (the same check as the ▶ Standard Disclosure but with a local police record check) to establish criminal backgrounds.

A ▶ Credit Check is an automated credit record search conducted through various major credit bureaus.

Introduction

Background checks became important to ▶ law enforcement about the time that large numbers of people started moving to cities as a by-product of the Industrial Revolution. Prior to that few people ever ventured far form their birthplace – a place where they were known and their history was known. The need to link people to their criminal histories drove police forces in London, Paris, and Buenos Aires to examine ▶ identification methodologies such as ▶ fingerprint recognition and ▶ anthropometry in the late 1800s – the surviving approaches are now classified as members of the science called biometrics. Simon Cole's 2001 book provides a good history of criminal identification [1].

In the post World War II era international travel became far more common than before the war. A parallel can be drawn with the movement during the Industrial Age within countries – now criminals and terrorists were freely crossing borders – hoping to leave their criminal/terrorist records behind. Even if the world's police records were all suddenly accessible over the Internet – they would not all be in the same character set. While many are in the Roman alphabet, others are in Cyrillic, Chinese characters, etc., this poses a problem for text search engines and investigators. Fortunately biometrics samples are insensitive to the nationality or country of origin of a person. Thus a search can be theoretically performed across the world using fingerprints or other enrolled biometric modalities. Unfortunately the connectivity of systems does not support such searches other than on a limited basis – through Interpol. If the capability to search globally were there the responses would still be textual and not necessarily directly useful to the requestor. One response to the challenge of international travel has been that nations collect biometric samples, such as the United Arab Emirates does with ▶ Iris Recognition, Overview, at their points of entry to determine if a person previously deported or turned down for a visa is attempting to reenter the country illegally.

A wide variety of positions of trust in both the public and private sectors require ▶ verification of suitability either as a matter of law or corporate policy. A person is considered suitable if the search for background impediments is negative. A position of trust can range from a police officer or teacher; to a new corporate employee who will have access to proprietary information and possibly a business's monetary assets; or to an applicant for a large loan or mortgage. Certain classes of jobs are covered by federal and state/provincial laws such as members of the military and school teachers/staff.

A background check is the process of finding information about someone which may not be readily available. The most common way of conducting a background check is to look up official and commercial records about a person. The need for a background check commonly arose when someone had to be hired for high-trust jobs such as security or in banking. Background checks while providing informed and less-subjective evaluations, however, also brought along their own risks and uncertainties.

Background checks require the "checking" party to collect as much information about the subject of the background check as is reasonably possible at the beginning of the process. Usually the subject completes a personal history form and some official document (e.g., a driver's license) is presented and photocopied. The information is used to increase the likelihood of narrowing the search to include the subject, but not too many others, with the same name or other attributes such as the same date and place of birth.

These searches are typically based on not only an individual's name, but also on other personal identifiers such as nationality, gender, place and date of birth, race, street address, driver's license number, telephone number, and Social Security Number. Without knowing where a subject really has lived it is very hard for an investigation to be successful without broad access to nationally aggregated records. There are companies that collect and aggregate these records as a commercial venture.

It is important to understand that short of some biometric sample (e.g., fingerprints) the collected information is not necessarily unique to a particular individual. It is well known that name checks, even with additional facts such as height, weight, and DOB can have varying degrees of accuracy because of identical or similar names and other identifiers. Reduced accuracy also results from clerical errors such as

misspellings, or deliberately inaccurate information provided by search subjects trying to avoid being linked to any prior criminal record or poor financial history.

In the US much of the required background information to be searched is publicly available but not necessarily available in a centralized location. Privacy laws limit access in some jurisdictions. Typically all arrest records, other than for juveniles, are public records at the police and courthouse level. When aggregated at the state level some states protect them while others sell access to these records. Other relevant records such as sex offender registries are posted on the Internet.

For more secure positions in the US, background checks include a "National Agency Check." These checks were first established in the 1950s and include a name-based search of FBI criminal, investigative, administrative, personnel, and general files. The FBI has a National Name Check Program that supports these checks. The FBI web site [2] provides a good synopsis of the Program:

Mission: The National Name Check Program's (NNCP's) mission is to disseminate information from FBI files in response to name check requests received from federal agencies including internal offices within the FBI; components within the legislative, judicial, and executive branches of the federal government; foreign police and intelligence agencies; and state and local law enforcement agencies within the criminal justice system.

Purpose: The NNCP has its genesis in Executive Order 10450, issued during the Eisenhower Administration. This executive order addresses personnel security issues, and mandated National Agency Checks (NACs) as part of the preemployment vetting and background investigation process. The FBI is a primary NAC conducted on all U.S. government employees. Since 11 September, name check requests have grown, with more and more customers seeking background information from FBI files on individuals before bestowing a privilege – whether that privilege is government employment or an appointment, a security clearance, attendance at a White House function, a Green card or naturalization, admission to the bar, or a visa for the privilege of visiting our homeland. . . .

Function: The employees of the NNCP review and analyze potential identifiable documents to determine whether a specific individual has been the subject of or has been mentioned in any FBI investigation(s), and if so, what (if any) relevant information may be disseminated to the requesting agency. It is important to note that the FBI does not adjudicate the final outcome; it just reports the results to the requesting agency.

Major Contributing Agencies: The FBI's NNCP Section provides services to more than 70 federal, state, and local governments and entities. . . . The following are the major contributing agencies to the NNCP:

- U.S. Citizenship and Immigration Services – Submits name check requests on individuals applying for the following benefits: asylum, adjustment of status to legal permanent resident, naturalization, and waivers.
- Office of Personnel Management – Submits name check requests in order to determine an individual's suitability and eligibility in seeking employment with the federal government.
- Department of State – Submits FBI name check requests on individuals applying for visas. . . . [2]

In the US government background checking process, a ▶ credit check "*is included in most background investigations except the basic NACI investigation required of employees entering Non-Sensitive (Level 1) positions* [3]."

Background checks were once the province of governments. Now commercial companies provide these services to the public, industry, and even to governments. These commercial checks rely on purchased, copied, and voluntarily submitted data from second and third parties. There are many commercial companies that accumulate files of financial, criminal, real estate, motor vehicle, travel, and other transactions. The larger companies spend substantial amounts of money collecting, collating, analyzing, and selling this information.

At the entry level, customers of these aggregators include persons "checking out" their potential roommates, baby sitters, etc. At the mid-level employers use these services to prescreen employees. At the high end the data is mined to target individuals for commercial and security purposes based on their background (e.g., financial and travel records.) The profiling of persons based on background information is disturbing in that the files are not necessarily accurate and rarely have biometric identifiers to identify people positively.

For an example of the problem, there is no need to look further than the news stories about post 9–11 name-based screening that kept Senator Kennedy on the no-fly list because he shared a name with a suspected person – and that was a government maintained file.

Historically the challenge in background checking has been (1) when people usurp another person's identity that "checks out" as excellent, (2) when people make up an identity and it is only checked for negative records not for its basic veracity, and (3) when persons try to hide their past or create a new past using multiple identities to gain benefits or privileges they might otherwise not be entitled to receive. A second identity could be created by simply changing their date and place of birth, of course it would not have much "depth" in that a simple check would reveal no credit history, no driver's license, etc. yet for some applicants checking is only to determine if the claimed identity has a negative history or not – not to see if the person really exists. All of these challenges render many name or number-based (e.g., social security number) background checks ineffective.

Several countries, states, and provinces are undertaking one relatively simple solution to stolen identities. As more and more records become digital, governments can link birth and death records – so a person cannot claim to be a person who died at a very early age and thus having no chance of a negative record. People were able to use these stolen identities as seeds for a full set of identification documents. Governments and financial institutions are also requiring simple proof of documented residence such as mail delivered to an applicant from a commercial establishment to the claimed address and a pay slip from an employer. Denying people easy ways to shift identities is a critical step in making background checks more reliable.

The most successful way to deal with these challenges has been to link persons with their positive (e.g., driver's license with a clean record) and negative history (e.g., arrest cycles) biometrically. The primary systems where this linkage is being done are in the provision of government services (motor vehicle administration and benefits management) and the criminal justice information arena (arrest records and court dispositions). Currently, few if any financial records are linked to biometric identifiers and the major information aggregators do not yet have biometric engines searching through the millions of records they aggregate weekly. The real reason they have not yet invested in this technology stems primarily from the almost total lack of access to biometric records other than facial images. This provides some degree of privacy for individuals while forcing credit bureaus to rely on linked textual data such as a name and phone number, billed to the same address as on file with records from a telephone company, with an employment record.

The inadequacy of name-based checks was redocumented in FBI testimony in 2003, regarding checking names of persons applying for Visas to visit the US. *Approximately 85% of name checks are electronically returned as having "No Record" within 72 hours. A "No Record" indicates that the FBI's Central Records System contains no identifiable information regarding this individual. . ."* This response does not ensure that the applicants are using their true identity but only that the claimed identity was searched against text-based FBI records – without any negative results.

The FBI also maintains a centralized index of criminal arrests, convictions, and other dispositions. The data is primarily submitted voluntarily by the states and owned by the states – thus limiting its use and dissemination. The majority of the 100 million plus indexed files are linked to specific individuals through fingerprints. The following information about the system is from a Department of Justice document available on the Internet [4].

This system is an automated index maintained by the FBI which includes names and personal identification information relating to individuals who have been arrested or indicted for a serious or significant criminal offense anywhere in the country. The index is available to law enforcement and criminal justice agencies throughout the country and enables them to determine very quickly whether particular persons may have prior criminal records and, if so, to obtain the records from the state or federal databases where they are maintained. Three name checks may be made for criminal justice purposes, such as police investigations, prosecutor decisions and judicial sentencing. In addition, three requests may be made for authorized noncriminal justice purposes, such as public employment, occupational licensing and the issuance of security clearances, where positive fingerprint identification of subjects has been made.

Name check errors are of two general types: (1) inaccurate or wrong identifications, often called "false positives," which occur when all three name

checks of an applicant does not clear (i.e., it produces one or more possible candidates) and the applicant's fingerprint search does clear (i.e., applicant has no FBI criminal record); and (2) missed identifications, often called "false negatives," which occur when the three name checks of an applicant's III clears (i.e., produces no possible candidates) and the applicant's fingerprint search does not clear (i.e., applicant has an FBI criminal record). Although errors of both types are thought to occur with significant frequency – based on the experience of state record repository and FBI personnel – at the time when this study was begun, there were no known studies or analyses documenting the frequency of such errors.

In contrast, fingerprint searches are based on a biometric method of identification. The fingerprint patterns of individuals are unique characteristics that are not subject to alteration. Identifications based on fingerprints are highly accurate, particularly those produced by automated fingerprint identification system (AFIS) equipment, which is in widespread and increasing use throughout the country. Analyses have shown that AFIS search results are 94–98% accurate when searching good quality fingerprints.

Because of the inaccuracies of name checks as compared to fingerprint searches, the FBI and some of the state criminal record repositories do not permit name-check access to their criminal history record databases for noncriminal justice purposes.

Where Do Biometrics Fit In?

When executing a background check there are several possible ways that biometric data can be employed. As seen governments can collect large samples (e.g., all ten fingers) to search large criminal history repositories. The large sample is required to ensure the search is cost effective and accurate. The time to collect all these fingerprints and extract the features can be measured in minutes, possibly more than ten, while the search time must be measured in seconds to deal with the national workloads at the central site.

Other programs such as driver's license applicant background checks are sometimes run using a single facial image. These are smaller files than fingerprints, collected faster using less costly technology, but have somewhat lower accuracy levels thus requiring more adjudication by the motor vehicle administrators.

As companies (e.g., credit card issuers) start to employ biometrics for convenience or brand loyalty – they are very likely to use the biometrics not just for identity verification at the point of sale but to weed out applicants already "blacklisted" by the issuer. These biometric samples will need to be of sufficient density to permit identification searches and yet have a subset that is "light weight" enough to be used for verification in less than a second at a point of sale.

Temporal Value of Background Checks

In the US under the best conditions a vetted person will have "passed a background check" to include an FBI fingerprint search, a NAC, a financial audit, personal interviews, and door-to-door field investigation to verify claimed personal history and to uncover any concerns local police and neighbors might have had. This is how the FBI and other special US agencies and departments check their applicants. Unfortunately, this is not sufficient. Robert Hanssen, Special Agent of the FBI was arrested and charged with treason in 2001 after 15 years of undetected treason and over 20 years of vetted employment.

Even more disturbing is a 2007 case where Nada Nadim Prouty pleaded guilty to numerous federal charges including unlawfully searching the FBI's Automated Case Support computer system. Ms. Prouty was hired by the FBI in 1999 and underwent a full background check that included fingerprints. In 2003 she changed employers, joining the CIA where she underwent some level of background check. Neither of these checks nor earlier checks by the then INS disclosed her having paid an unemployed American to marry her to gain citizenship.

Without being caught, criminals have the same clean record as everyone else, with or without biometrics being used in a background check. While FBI agent's fingerprints are kept in the FBI's AFIS system, those of school teachers and street cops are not. This means if any of them were arrested only the FBI's employees' fingerprints would lead to notification of the subject's employer. Rap (allegedly short for Record of Arrests and Prosecutions) sheets are normally provided in response to fingerprint searches. A relatively new process called Rap-Back permits agencies requesting a background check to enroll the fingerprints such that if there is a subsequent arrest

the employer will be notified. Rap-Back and routine reinvestigations addresses part of the temporal problem but in the end there is no guarantee that a clean record is not a misleading sign – just an indicator of no arrests, which is not always a sign of trustworthiness.

Privacy Aspects

Performing a complete and accurate background check can cause a conflict with widely supported privacy laws and practices. The conflicts come from most "privacy rights" laws being written to inhibit certain government actions not to uniformly limit commercial aggregation and sharing of even questionable data on a background-data-for-fee basis.

Robert O'Harrow's book [5], points out two serious flaws on the privacy side of the commercial background check process.

- "Most employees who steal do not end up in public criminal records. Dishonest employees have learned to experience little or no consequences for their actions, especially in light of the current tight labor market," ChoicePoint tells interested retailers. "A low-cost program is needed so companies can afford to screen all new employees against a national theft database." The database works as a sort of blacklist of people who have been accused or convicted of shoplifting [6].
- Among other things the law restricted the government from building databases of dossiers unless the information about individuals was directly relevant to an agency's mission. Of course, that's precisely what ChoicePoint, LexisNexis, and other services do for the government. By outsourcing the collection of record, the government doesn't have to ensure the data is accurate, or have any provisions to correct it in the same way it would under the Privacy Act [7].

These companies have substantially more information on Americans than the government. O'Harrow reports that Choice Point has data holdings of an unthinkable size:

- Almost a billion records added from Trans Union twice a year
- Updated phone records (numbers and payment histories) from phone companies – for over 130 million persons
- A Comprehensive Loss Underwriting Exchange with over 200 million claims recorded
- About 100 million criminal records
- Copies of 17 billion public records (such as home sales and bankruptcy records)

The United Nations International Labor Organization (ILO) in 1988 described "*indirect discrimination*" *as occurring when an apparently neutral condition, required of everyone, has a disproportionately harsh impact on a person with an attribute such as a criminal record.*" [8] Thus pointing out the danger of cases where criminal records "*include charges which were not proven, investigations, findings of guilt with non-conviction and convictions which were later quashed or pardoned. It also includes imputed criminal record. For example, if a person is denied a job because the employer thinks that they have a criminal record, even if this is not the case* [9]."

This problem is recognized by the Australian government, which quotes the above ILO words in its handbook for employees. The handbook goes on to say, "*The CRB recognises that the Standard and Enhanced Disclosure information can be extremely sensitive and personal, therefore it has published a Code of Practice and employers' guidance for recipients of Disclosures to ensure they are handled fairly and used properly*" [10].

Applications

Background checks are used for preemployment screening, establishment of credit, for issuance of Visas, as part of arrest processes, in sentencing decisions, and in granting clearances. When a biometric check is included, such as a fingerprint-based criminal-records search, there can be a higher degree of confidence in the completeness and accuracy of that portion of the search.

Seemingly secure identification documents such as biometric passports do not imply a background check of the suitability of the bearer – but only that the person it was issued to is a citizen of the issuing country. These documents permit positive matching of the bearer to the person the document was issued to – identity establishment and subsequent verification of the bearer.

It is unfortunately easy to confuse the two concepts: a clean background and an established identity. The US Government's new Personal Identity Verification (PIV) card, on the other hand, implies both a positive

background check and identity establishment. The positive background check is performed through a fingerprint-based records search. The identity is established when a facial image, name, and other identity attributes are locked to a set of fingerprints. The fingerprints are then digitally encoded and loaded on the PIV smart card; permitting verification of the bearer's enrolled identity at a later date, time, and place.

Summary

Background checks are a necessary but flawed part of the modern world. Their importance has increased substantially since the terrorist attacks on 9/11 in the US, 11-M in Spain, and 7/7 in London. Governments use them within privacy bounds set by legislatures but seem to cross into a less constrained world when they use commercial aggregators. Industry uses them in innumerable process – often with little recourse by impacted customers, employees, and applicants. Legislators are addressing this issue but technology is making the challenge more ubiquitous and at an accelerating rate.

Biometric attributes linked to records reduce the likelihood of them being incorrectly linked to a wrong subject. This is a promise that biometrics offers us – yet the possible dangers in compromised biometric records or systems containing biometric identifiers must be kept in mind.

Related Entries

- ▶ Fingerprint Matching, Automatic
- ▶ Fingerprint Recognition, Overview
- ▶ Fraud Reduction, Application
- ▶ Fraud Reduction, Overview
- ▶ Identification
- ▶ Identity Theft Reduction
- ▶ Iris Recognition at Airports and Border-Crossings
- ▶ Law Enforcement
- ▶ Verification

References

1. Simon A. Cole.: Suspect Identities: A History of Fingerprinting and Criminal Identification. Harvard University Press (2001), ISBN 0-6740-1002-7
2. http://www.fbi.gov/hq/nationalnamecheck.htm
3. HHS personnel Security/Suitability Handbook; SDD/ASMB 1/98
4. http://www.ojp.usdoj.gov/bjs/pub/pdf/iiince.pdf
5. Robert O'Harrow, Jr.: No Place To Hide. Free Press, New York (2005), ISBN 0-7432-5480-5
6. Robert O'Harrow, Jr.: No Place To Hide, p. 131. Free Press, New York (2005), ISBN 0-7432-5480-5
7. Robert O'Harrow, Jr.: No Place To Hide, p. 137. Free Press, New York (2005), ISBN 0-7432-5480-5
8. International Labour Conference, Equality in Employment and Occupation: General Survey by the Committee of Experts on the Application of Conventions and Recommendations ILO, Geneva (1988)
9. On the record; Guidelines for the prevention of discrimination in employment on the basis of criminal record; November 2005; Minor Revision September 2007; Australian Human Rights and Equal Opportunity Commission.
10. http://www.crb.gov.uk/Default.aspx?page = 310

Background Subtraction

Segmentation of pixels in a temporal sequence of images into two sets, a moving foreground and a static background, by subtracting the images from an estimated background image.

▶ Gait Recognition, Silhouette-Based

Back-of-Hand Vascular Pattern Recognition

▶ Back-of-Hand Vascular Recognition

Back-of-Hand Vascular Recognition

Alex Hwansoo Choi
Department of Information Engineering, Myongji University, Seoul, South Korea

Synonyms

Back-of-hand vascular pattern recognition; Hand vascular recognition; Hand vein identification, Hand vein verification

Definition

The back-of-hand vascular recognition is the process of verifying the identity of individuals based on their subcutaneous vascular network on the back of the hand. According to large-scale experiments, the pattern of blood vessels is unique to each individual, even among identical twins; thereby the pattern of the hand blood vessels on the back of the hand can be used as distinctive features for verifying the identity of individuals. A simple back-of-hand vascular recognition system is operated by using ▶ near-infrared light to illuminate on the back of the hand. The deoxidized hemoglobin in blood vessels absorb more infrared rays than surrounding tissues and cause the blood vessels to appear as black patterns in the resulting image captured by a camera, sensitive to near-infrared illumination. The image of back-of-hand vascular patterns is then pre-processed and compared with the previously recorded vascular pattern ▶ templates in the database to ▶ verify the identity of the individual.

Introduction

▶ Biometric recognition is considered as one of the most advanced security method for many security applications. Several biometric technologies such as fingerprint, face, and hand geometry have been researched and developed in recent years [1]. Compared with traditional security methods such as pass codes, passwords, or smart cards, the biometric security schemes show many priority features such as high level security and user convenience. Therefore, biometric recognition systems are being widely deployed in many different applications.

The back-of-hand vascular pattern is a relatively new biometric feature containing complex and stable blood vessel network that can be used to discriminate a person from the other. The back-of-hand vascular pattern technology began to be considered as a potential biometric technology in the security field in early 1990s. During this period, the technology became one of the most interesting topics in biometric research community that received significant attention. One of the first paper to bring this technology into discussion was published by Cross and Smith in 1995 [2]. The paper introduced the ▶ thermographic imaging technology for acquiring the subcutaneous vascular network of the back of the hand for biometric application. However, the thermographic imaging technology is strongly affected by temperature from external environment; therefore, it is not suitable to apply this technology to general out-door applications.

The use of back-of-hand vascular recognition in general applications became possible when new imaging techniques using near-infrared illumination and low-cost camera have been invented [3]. Instead of using far-infrared light and thermographic imaging technology, this technology utilizes the near-infrared light to illuminate the back of the hand. Due to the difference in absorption rate of infrared radiation, the blood vessels would appear as black patterns in the resulting image. The cameras to photograph the back-of-hand vascular pattern image can be any low-cost cameras that are sensitive to the range of near-infrared light.

Although the back-of-hand vascular pattern technology is still an ongoing area of biometric research, it has become a promising identification technology in biometric applications. A large number of units deployed in many security applications such as information access control, homeland security, and computer security provide evidence to the rapid growth of the back-of-hand vascular pattern technology. Compared to the other existing biometric technologies, back-of-hand vascular pattern technology has many advantages such as higher authentication accuracy and better ▶ usability. Thereby, it is suitable for the applications in which high level of security is required. Moreover, since the back-of-hand vascular patterns lies underneath the skin, it is extremely difficult to spoof or steal. In addition, lying under skin surface, back-of-hand vascular pattern remains unaffected by inferior environments. Therefore, the back-of-hand vascular pattern technology can be used in various inferior environments such as factories, army, and construction sites where other biometric technologies have many limitations. Because of these advanced features, the back-of-hand vascular pattern technology is used in public places.

Development History of the Back-of-Hand Vascular Recognition

As a new biometric technology, back-of-hand vascular pattern recognition began to receive the attention

from biometric community from 1990s. However, the launch of the back-of-hand vascular recognition system into the market was first considered from 1997. The product model named BK-100 was announced by BK Systems in Korea. This product was sold mainly in the local and Japanese market. In the early stage of introduction, the product had limitations for physical access control applications. More than 200 units have been installed in many access control points with time and attendance systems in both Korea and Japan. Figure 1(a) shows a prototype of the BK-100 hand vascular recognition system.

The first patent on the use of the back-of-hand vascular pattern technology for personal identification was published in 1998 and detailed in reference [4]. The invention described and claimed an apparatus and method for identifying individuals through their subcutaneous hand vascular patterns. Consecutively, other subsequent commercial versions, BK-200 and BK-300, have been launched in the market. During the short period of time from the first introduction, these products have been deployed in many physical access control applications.

The technology of back-of-hand vascular pattern recognition was continuously enhanced and developed by many organizations and research groups [5–11]. However, one of the organizations that made promising contributions to the development of the back-of-hand recognition technology is Techsphere Co., Ltd. in Korea. As the results from these efforts, a new commercial product under the name VP-II has been released. Many advanced digital processing technologies have been applied to this product to make it a reliable and cost-effective device. With the introduction of the new product, the scanner became more compact to make the product suitable to be integrated in various applications. The product also provided better user interface to satisfy user-friendly requirements and make the system highly configurable. Figure 1(b) shows a prototype of the VP-II product.

Back-of-Hand Vascular Recognition. Figure 1 Prototype of hand vascular recognition system; **(a)** BK-100 and **(b)** VP-II product.

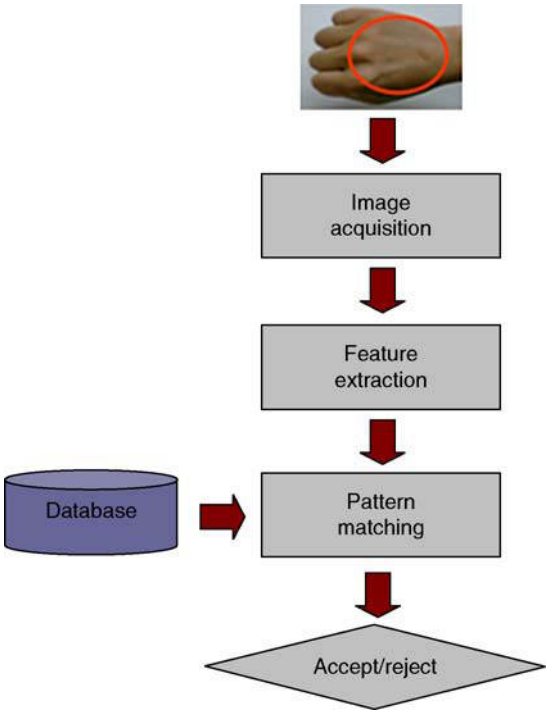
Various organizations and research groups are spending efforts to develop and enhance the back-of-hand vascular pattern technology. Thousands of back-of-hand products have been rapidly installed and successfully used in various applications. Researches and product enhancements are being conducted to bring more improvements to products. Widespread international attention from biometric community will make the back-of-hand vascular pattern technology as one of the most promising technologies in security field.

Underlying Technology of Back-of-Hand Vascular Recognition

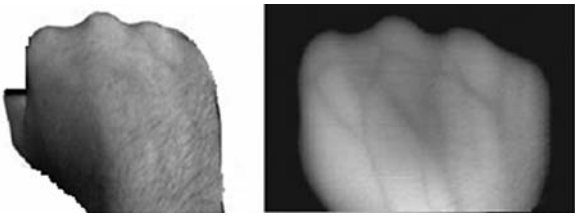
To understand the underlying technology of the back-of-hand vascular recognition, the operation of a typical back-of-hand vascular recognition system should be considered. Similar to other biometric recognition system, the back-of-hand vascular recognition system often composes of different modules including image acquisition, feature extraction, and pattern matching. Figure 2 shows a typical operation of the back-of-hand vascular recognition system.

Image Acquisition

Since the back-of-hand vascular pattern lies underneath the skin, it cannot be seen by the human eye. Therefore, it cannot use the visible light that occupies a very narrow band (approx. 400–700 nm wavelength) for photographing the back-of-hand vascular patterns. The back-of-hand vascular pattern image can be captured under the near-infrared light (approx. 800–1000 nm wavelength). The near-infrared light can penetrate into the human tissues to approximately 3 mm depth [10]. The blood vessels absorb more infrared radiation than the surrounding tissues and appear as black patterns in the resulting image. The camera used to capture the image of back-of-hand vascular pattern can be any low-cost camera that is sensitive to the range of near-infrared light. Figure 3 shows an example of images obtained by visible light and near-infrared light.



Back-of-Hand Vascular Recognition. **Figure 2** Operation of a typical back-of-hand vascular recognition system.



Back-of-Hand Vascular Recognition. **Figure 3** The back-of-hand images obtained by visible light (left) and by infrared light (right).

Pattern Extraction

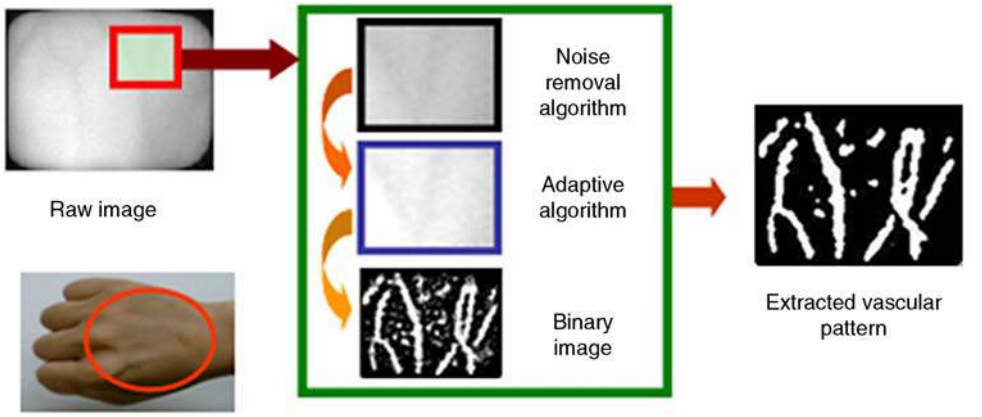
One of the important issues in the back-of-hand vascular recognition is to extract the back-of-hand vascular pattern that can be used to distinguish an individual from the others. Pattern extraction module is to accurately extract the back-of-hand vascular patterns from raw images which may contain the undesired noises and other irregular effects. The performance of the back-of-hand vascular recognition system strongly depends on the effectiveness of the pattern extraction module. Therefore, the pattern extraction module often consists of various advanced image processing algorithms to remove the noises and irregular effects, enhance the clarity of vascular patterns, and separate the vascular patterns from the background. The final vascular patterns obtained by the pattern extraction algorithm are represented as binary images. Figure 4 shows the procedure of a typical feature extraction algorithm for extracting back-of-hand vascular patterns from raw images. After the pattern extraction process, there still could be salt-and-pepper type noises. Thus, noise removal filters such as medial filters may be applied as a post-processing step.

Pattern Matching

The operation of a back-of-hand vascular recognition system is based on comparing the back-of-hand vascular pattern of a user being authenticated against pre-registered back-of-hand patterns stored in the database. The comparison step is often performed by using different type of pattern matching algorithms to generate a matching score. The structured matching algorithm is utilized if the vascular patterns are represented by collections of some feature points such as line-endings and bifurcations [13]. If the vascular patterns are represented by binary images, the template matching algorithm is also utilized [14]. The matching score is then used to compare with the pre-defined system threshold value to decide whether the user can be authenticated. For more specific performance figures for each algorithm, readers are referred to [4–7].

Applications of Back-of-Hand Vascular Recognition

The ability to verify identity of individuals has become increasingly important in many areas of modern life, such as electronic governance, medical administration systems, access control systems for secured areas, and passenger ticketing, etc. With many advanced features such as high level of security, excellent usability, and difficulty in spoofing, the back-of-hand vascular recognition systems have been deployed in a wide range of practical applications. The practical applications of the back-of-hand vascular recognition systems can be summarized as following:



Back-of-Hand Vascular Recognition. **Figure 4** The flow chart of a typical feature extraction algorithm.

Office access control and Time Attendance: The wide use of back-of-hand vascular recognition technology is physical access control and identity management for time and attendance. The recognition systems utilizing the back-of-hand vascular technology are often installed to restrict the access of unauthorized people. The integrated applications with back-of-hand vascular recognition systems will automatically record the time of entering and leaving the office for each employee. Furthermore, the time and attendance record for each employee can be automatically fed to the resource management program of the organization. This provides a very effective and efficient way to manage the attendance and over-time payment at large-scale organizations.

Port access control: Due to the overwhelming security climate in recent years and fear of terrorism, there has been a surge in demand for accurate biometric authentication methods to establish a security fence in many ports. Airports and seaports are the key areas through which terrorists may infiltrate. Due to its high accuracy and usability, fast recognition speed, and user convenience, the back-of-hand vascular recognition systems are being employed for access control in many seaports and airports. For example, back-of-hand vascular recognition systems are being used in many places at Incheon International Airport and many airports in Japan [14]. In addition, major Canadian seaports (Vancouver and Halifax) are fully access-controlled by back-of-hand systems.

Factories and Construction Sites: Unlike other biometric features which can be easily affected by dirt or oil, the back-of-hand vascular patterns are not easily disturbed because the features lie under the skin of human body. Therefore, the back-of-hand vascular pattern technology is well accepted in applications exposed to inferior environments such as factories or construction sites. The strengths and benefits of the back-of-hand vascular pattern technology become more obvious when it is used in these applications because other existing biometric technologies show relatively low usability and many limitations when used in inferior environments.

Summary

The back-of-hand vascular pattern technology has been researched and developed in the recent decades. In a relatively short period, it has gained considerable attention from biometric community. The rapidly growing interest in the back-of-hand vascular pattern technology is confirmed by the large number of research attempts which have been conducted to improve the technology in recent years. Although the back-of-hand vascular pattern has provided a higher accuracy and better usability in comparison with other existing biometric technologies, more research need to be performed to make it more robust and tolerant technology in various production conditions. The future research should focus on development of higher quality image capture devices, advanced feature extraction algorithms, and more reliable pattern matching algorithms to resolve pattern distortion issue.

Related Entries

- ▶ Finger Vein
- ▶ Finger Vein Biometric Algorithm
- ▶ Finger Vein Pattern Imaging
- ▶ Finger Vein Reader
- ▶ Palm Vein Image Device
- ▶ Palm Vein

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An Introduction to Biometric. IEEE Trans. Circ. Syst. Video Technol. **14**(1), 4–20 (2004)
2. Cross, J.M., Smith, C.L.: Thermographic Imaging of the Subcutaneous Vascular Network of the Back of the Hand for Biometric Identification. In: Proceedings of the IEEE Annual Conference, pp. 20–35. October (1995)
3. Im, S.K., Park, H.M., Kim, Y.W., Han, S.C., Kim, S.W., Kang, C.H.: Biometric Identification System by Extracting Hand Vein Patterns. J. Korean Phys. Soc. **38**(3), 268–272 (2001)
4. Choi, H.S., Systems, B.K.: Apparatus and Method for Identifying Individuals Through Their Subcutaneous Vein Patterns and Integrated System using Said Apparatus and Method. US Patent, no.6301375, (2001)
5. Im, S.K., Choi, H.S., Kim, S.W.: Design for an Application Specific Processor to Implement a Filter Bank Algorithm for Hand Vascular Pattern Verification, J. Korean Phys. Soc. **41**, 461–467 (2002)
6. Im, S.K., Choi, H.S.: A Filter Bank Algorithm for Hand Vascular Pattern Biometrics. In: Proceedings of ICCARV'02, pp. 776–781 (2002)
7. Im, S.K., Choi, H.S., Kim, S.W.: A Direction-based Vascular Pattern Extraction Algorithm for Hand Vascular Pattern Verification. ETRI J. **25**(2), 101–108 (2003)
8. Im, S.K., Park, H.M., Kim, S.W., Chung, C.K., Choi, H.S.: Improved Vein Pattern Extracting Algorithm and Its Implementation. In: ICCE 2000, pp. 2–3, June (2000)
9. TanakaT., Kubo: N.: Biometric Authentication by Hand Vein Patterns. In: SICE 2004, vol. 1, pp. 249–253, August (2004)
10. Ding, Y., Zhuang, D., Wang, K.: A Study of Hand Vein Recognition Method. In: Proceedings of the IEEE International conference on Mechatronics & Automation, pp. 2106–2110. July (2005)
11. Badawi, A.M.: Hand Vein Biometric Verification Prototype: A Testing Performance and Patterns Similarity. In: International Conference on Image Processing, Computer Vision, and Pattern Recognition, pp. 3–9 (2006)
12. Wang, L., Leedham, G.: Near- and Far- Infrared Imaging for Vein Pattern Biometrics. In: Proceedings of the IEEE International Conference on Video and Signal, pp. 52–59 (2006)
13. Kumar, A., Prathyusha, K.V.: Personal authentication using hand vein triangulation. In: Proceedings SPIE Conf Biometric Technology for human identification, vol. 6944, Orlando, pp. 69440E-69440E-13, Mar. (2008)
14. Choi, A.H., Tran, C.N.: Handbook of Biometrics: Hand vascular pattern recognition technology. Springer, New York (2008)

Barefoot Morphology Comparison

This describes the comparison carried out by a forensic expert to determine if two footprints could, or could not, have been made by the same person.

- ▶ Forensic Barefoot Comparisons

Base Classifier

This term is used to indicate the base component of a multiple classifier system. In other words, a multiple classifier system is made up by a set of base classifiers. Some authors use this term only when the multiple classifier system is designed using a single classification model (e.g., a decision tree) and multiple versions of this base classifier are generated to build the multiple classifier system.

- ▶ Multiple Classifier Systems

Baseline Algorithm

Baseline algorithm is a simple, yet reasonable, algorithm that is used to establish minimum expected performance on a dataset. For instance, the eigenfaces approach based on principal component analysis is the baseline algorithm for face recognition. And, the silhouette correlation approach establishes the baseline for gait recognition.

- ▶ Performance Evaluation, Overview

Baum-Welch Algorithm

The Baum-Welch algorithm is the conventional, recursive, efficient way to estimate a Hidden Markov Model,

that is, to adjust the parameters of the model given the observation sequence. The solution to this problem permits to develop a method to train self-learning classifiers.

► Hidden Markov Models

Bayes Decision Theory

A probabilistic framework for assigning an input pattern (e.g., a feature vector) to a class (or category) so as to minimize the risk associated with misclassification. The risk itself is computed as a function of several factors including the conditional probabilities describing the likelihood that the input pattern belongs to a particular class and the cost of misclassification as assessed by the practitioner. In some cases, the risk function is defined purely in terms of the probability of error. The various probabilities characterizing the framework are estimated using a set of training data comprising patterns whose class information is known beforehand.

► Fusion, Score-Level

Bayes Rule

Bayes theorem or Bayes rule allows the estimation of the probability that, a hypothesis H is true when presented with a set of observations or evidence E. Let $P(H)$ be the best estimate of the probability that hypothesis H is true prior to the availability of evidence E. Hence, $P(H)$ is known as the *prior probability* of H. Let $P(E|H)$ be the conditional probability (*likelihood*) of observing the evidence E given that, H is true and $P(E)$ be the marginal probability of E. Then, the *posterior probability* of hypothesis H given evidence E is.

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

According to the Bayes rule, the posterior probability is proportional to the product of the likelihood and the prior probabilities.

► Soft Biometrics

Bayesian Approach/Likelihood Ratio Approach

This approach used for interpreting evidence is based on the mathematical theorem of Reverend Thomas Bayes, stating the posterior odds are equal to the prior odds multiplied the likelihood ratio (LR). If using prior odds, one will speak of the Bayesian approach, and if using only the LR, one will speak of the likelihood ratio approach.

In forensic science, the weight of evidence E (DNA, glass, fingerprints, etc.) is often assessed, using the ratio of two probabilities estimated given by two propositions (i.e., LR). One hypothesis is suggested by the prosecution (H_p) and the other by the defence (H_d). Two propositions could be for example "The blood recovered from the crime scene comes from the suspect" versus "The blood recovered from the crime scene does not come from the suspect, but from someone else, unrelated to him." The likelihood ratio (LR) is therefore constructed as the ratio of the two probabilities of the observations given in each proposition. It can take any value between zero and infinity. Values inferior to one favor the defence proposition and values above favor the prosecution proposition. This metric is used worldwide and has been the subject of numerous publications. Verbal scales for reporting LR have been suggested, as an example values from 1,000 and upwards would provide very strong support for H_p.

The likelihood ratio approach permits to evaluate evidence using a balanced, logical, and scientific view. It helps avoiding erroneous reasoning, such as the prosecution or the defence fallacy.

► Forensic DNA Evidence

Bayesian Hypothesis Test

Given a segment of speech Y and a speaker S, the speaker verification task consists indetermining if Y was spoken by S or not. This task is often stated as basic hypothesis test between two hypotheses: If Y comes from the hypothesized speaker S it is $H0$, and if Y is not from the hypothesized speaker S it is $H1$.

A likelihood ratio (LR) between these two hypotheses is estimated and compared to a decision threshold θ. The LR test is given by:

$$\mathrm{LR}(Y, H0, H1) = \frac{p(y|H0)}{p(y|H1)} \quad (1)$$

where Y is the observed speech segment, $p(Y|H0)$ is the likelihood function for the hypothesis $H0$ evaluated for Y, $p(Y|H1)$ is the likelihood function for $H1$, and θ is the decision threshold for accepting or rejecting $H0$. If $LR(Y, H0, H1) > \theta$, $H0$ is accepted else $H1$ is accepted.

A model denoted by λ_{hyp} represents H0, which is learned using an extract of speaker S voice. The model $\lambda_{\overline{\mathrm{hyp}}}$ represents the alternative hypothesis, H1, and is usually learned using data gathered from a large set of speakers. The likelihood ratio statistic becomes $\frac{p(Y|\lambda_{\mathrm{hyp}})}{p(Y|\lambda_{\overline{hyp}})}$. Often, the logarithm of this statistic is used giving the *log*LR (LLR):

$$\mathrm{LLR}(Y) = \log(p(Y|\lambda_{\mathrm{hyp}})) - (p(Y|\lambda_{\overline{\mathrm{hyp}}})). \quad (2)$$

▶ Speaker Matching

Behavioral Biometrics

Behavioral biometrics is the class of biometrics based on various human actions as opposed to physical characteristics. Typically, behavioral biometrics is used only in verification frameworks. Examples of behavioral biometrics include: keystroke recognition, speaker/voice recognition, and signature. Behavioral biometrics is arguably more replaceable than physiological biometrics, as the context in which they are based can often be changed (i.e., keystroke recognition, voice, etc.).

▶ Keystroke Recognition

BIAS

▶ Biometric Identity Assurance Services

Bias-Variance Decomposition

Bias-variance decomposition is an important tool for analyzing machine learning approaches. Given a learning target and the size of training data set, it breaks the expected error of a learning approach into the sum of three nonnegative quantities, i.e., the *intrinsic noise*, the *bias*, and the *variance*. The intrinsic noise is a lower bound on the expected error of any learning approach on the target; the bias measures how closely the average estimate of the learning approach is able to approximate the target; the variance measures how much the estimate of the learning approach fluctuates for the different training sets of the same size.

▶ Ensemble Learning

Bi-directional Reflectance Distribution Function (BRDF)

Let us assume that the irradiance received by an elementary surface from a point light source is ΔE, and that the radiance from the elementary surface emits in an outgoing direction toward the viewer is ΔL. BRDF is defined as the ratio of $\Delta E/\Delta L$, i.e., the ratio of the radiance in the outgoing direction to the incident irradiance. The unit of BRDF is sr^{-1}.

▶ Image Formation

Bifurcation

The point at which a blood vessel splits or forks into two branches. An individual's unique pattern of retinal blood vessel bifurcations can be used as a feature space for retinal biometric encoding.

▶ Simultaneous Capture of Iris and Retina for Recognition

Binary Hypothesis

The binary hypothesis represents a decision maker with only two hypotheses to choose from. For biometrics, this usually means that the sensor has identified the genuine user or an imposter.

► Fusion, Decision-Level

Binary Morphology

Operations on binary images using convolution-type kernels (the "structuring elements") and basic set operations and image translations. Also called image morphology or mathematical morphology.

► Segmentation of Off-Axis Iris Images

Binomial Distribution

A major class of discrete probability distribution that describes the likelihood of outcomes from runs of Bernoulli trials (conceptually coin tosses with two possible results from each toss, having stable but not necessarily equal probabilities). A binomial distribution is described by two parameters: the probability of one of the outcomes (which implies the probability of the other); and the number of trials (or coin tosses) conducted. If the two outcome probabilities are equal, then the distribution is symmetrical; otherwise it is not. If one measures the *fraction* of outcomes of one type that occur in a certain number of tosses, then the mean of the distribution equals the probability of that outcome, and its standard deviation varies inversely as the square-root of the number of trials conducted. Thus, the larger the number of trials, the tighter this distribution becomes. The tails of a binomial distribution attenuate very rapidly because of the factorial combinatorial terms generating it, particularly when the number of trials is large. The binomial should not be confused with, nor interchanged with, a Gaussian distribution, which describes continuous instead of discrete random variables, and its domain is infinite unlike the compact support of the binomial. Under certain statistical conditions, even correlated Bernoulli trials generate binomial distributions; the effect of correlation is to reduce the effective number of trials. When IrisCodes from different eyes are compared, the distribution of normalized fractional Hamming distance scores follows a binomial distribution, since comparisons of IrisCode bits are effectively Bernoulli trials.

► Score Normalization Rules in Iris Recognition

BioAPI

BioAPI is Biometric Application Programming Interface. BioAPI 2.0 is a widely recognized international standard created by the BioAPI consortium and defined in ISO/IEC 19784-1:2005.

► Biometric Technical Interface, Standardization
► Finger Vein Pattern Imaging

BioAPI Framework

A module supplied by one vendor that provides the linkage (via the BioAPI API and SPI interfaces) between application modules and Biometric Service Providers from other independent vendors that support the standardized interfaces.

► Biometric Technical Interface, Standardization

BioAPI Interworking Protocol

A fully specified protocol running over the Internet (including Secure Socket Layer) that allows one BioAPI

Framework to communicate with another BioAPI Framework. This allows applications in one computer system to communicate with Biometric Service Providers (BSPs) in one or more computer systems, where the applications and BSPs are only aware of the local BioAPI API or SPI interface, and are not aware of the communications protocol. It provides a seamless integration of application and BSP modules running in different systems to provide for many forms of distributed biometric application.

► Biometric Technical Interface, Standardization

Biological Motion

The motion arising from the movement of living things. Although the term is consistent with viewing an action in natural conditions with full lighting, it is commonly used in the field of visual perception to denote motion patterns arising from viewing conditions with reduced visual information such as just the motion of specific points on the body or the silhouette. Human gait is a common example of biological motion.

► Psychology of Gait and Action Recognition

Biometric Algorithms

Yi Chen[1], Jean-Christophe Fondeur[2]
[1]Department of Computer Science and Engineering, Michigan State University, MI, USA
[2]Sagem Sécurité, Paris, France

Synonym

Biometric Engines

Definition

Biometric algorithms are automated methods that enable a biometric system to recognize an individual by his or her anatomical/behavioral traits [1]. They consist of a sequence of automated operations performed by the system to verify or identify its ownership. These operations include quality assessment, enhancement, feature extraction, classification/indexing, matching and fusion, as well as compression algorithms, often used to reduce storage space and bandwidth.

Introduction

Biometric recognition is achieved by comparing the acquired biometric sample (the "query") with one or more biometric samples that have been captured previously and stored in the system database (the "reference" or "gallery"). The process of creating the database is called enrollment. The process of comparing samples is called verification if the query comes along with a claimed identity (in this case the "query" is compared to the biometric data of the claimed identity), or identification if no identity claim is made (in this case the "query" is compared to all the biometric data in the database).

The biometric sample is acquired by a biometric device and produces an electronic representation of high-dimensional signals (e.g., fingerprint or face images, signature dynamics) [2]. Most often, to avoid the "► curse of dimensionality," these high-dimensional signals are not directly compared; instead, a more compact representation of the signal – called "template" – is extracted from the raw signal and is used for the comparison. The various processes used to compare them are called biometric algorithms. These processes include assessing and enhancing the quality of the biometric signal, extracting and matching salient features, and information fusion at various stages. Compression and classification/indexing are also key components of biometric algorithms to optimize the resources needed (space and time).

Biometric techniques are effective to recognize people because the characteristics of biometric traits are distinct to each individual. In practice, however, variations (inherent in the biometric trait or how it is presented during acquisition) and noise, as well as intrinsic limitations of biometric sensing techniques can cause the accuracy of the system to drop significantly. It is necessary to develop biometric algorithms that are robust to these variations; namely, to extract salient and reproducible features from the input and to

match these features efficiently and effectively with the templates in the database. Addressing all the problems requires the combination of various techniques to obtain the optimal robustness, performance, and efficiency, which is a key step in biometric algorithm design.

Compression

Many applications require storage or transmission of the biometric data (e.g., images). These data can be large and it is often desirable to compress them to save storage space or transmission bandwidth. This compression can be either lossless or lossy. Lossless compression algorithms guarantee that every single bit of the original signal is unchanged after the data is uncompressed. Higher compression ratio can be achieved with lossy compression at the cost of altering the original signal. Artifacts introduced by lossy compression may interfere with subsequent feature extraction and degrade the matching results.

Biometric systems often use lossy compression, chosen in such a way that a minimal amount of critical information is lost during the compression, to achieve the best balance between data quality and representation size. Standardization bodies have defined compression protocols for each biometric so that any user of the system can reconstruct the original signal. They also specify the compression ratio that must be used to preserve the quality of the biometric data. As an example, standards currently exist for the compression of fingerprints (WSQ for 500ppi and JPEG-2,000 for 1,000ppi), facial images (JPEG-2,000), voice data (CELP) [3, 4].

Quality Assessment

Biometric quality refers to the usefulness of a biometric sample in terms of the amount of discriminatory information. Quality assessment is the algorithm that calculates and assigns a quantitative quality score to a biometric sample based on its character (e.g., inherent features), fidelity (e.g., signal to noise ratio), or utility (e.g., correlation with system performance) [5].

Quality measure can be used for various applications in a biometric system: (1) to provide quality feedback upon enrollment to improve the operational efficiency of biometric systems; (2) to improve the matching performance of biometric systems, e.g., local quality can be used to assist feature extraction and assign confidence to features during matching; and (3) to improve performance of multi-biometric systems, e.g., quality can be used to derive weights or statistical significance of individual sample or modality in fusion.

There are two main paradigms for quality assessment algorithms: a "bottom-up" approach reflecting character and fidelity; and a "top-down" approach based on observed utility [5]. In the "bottom-up" approach, quality measure is used to determine a sample's "improvability" (i.e., the improvement that can be gained by recapturing the biometric). If a sample does not inherently have many features, recapturing will not benefit the performance. On the other hand, if the signal to noise ratio is very high, recapturing may help obtain additional salient features. In the "top-down" approach, the utility of a sample is used to determine a performance estimate. This estimate can be used to disregard (emphasize) features that have strong (weak) correlation with utility.

Development of quality assessment algorithms and algorithms that use the estimated quality information is an active area of research in biometric community. The NIST biometric quality workshop [6] provides a forum for the community to share new research and development in biometric quality assessment. An open source software to measure fingerprint quality has also been developed and released by NIST [7]. Standards committees from around the world are working to incorporate the concept of quality into the biometric standards, e.g., ISO/IEC 29794 [8], with the aim of uniform interpretation and interoperability of quality scores.

Enhancement

Enhancement, in the context of biometrics, is the process of improving the signal quality with or without knowing the source of degradation (this definition includes restoration). The general goal is to increase the signal to noise ratio, although, many interpretations of signal/noise can be applicable. Enhancement typically employs prior knowledge about the acquired signal to facilitate automatic feature extraction algorithms or to provide better visualization for manual processing.

The quality of a signal can be affected by environmental conditions, sensor noise, uncooperative/untrained subjects, inherent low quality biometrics, etc. In order to ensure that the performance of a biometric algorithm will be robust with respect to the quality of the acquired signal, additional algorithms/heuristics must be employed to improve the clarity of the desired traits in the signal. Different types of normalization (e.g., histogram equalization) or filtering approaches (e.g., Gabor wavelets) can be employed to separate noise from biometric signals [9]. Segmentation (i.e., detecting the meaningful part of the signal and discarding the background) is another example of enhancement that is classically used.

Feature Extraction

During feature extraction, the biometric data is processed to extract a set of salient and discriminatory features that represent the underlying biometric trait. These features can either have a direct physical counterpart (e.g., minutiae for fingerprints), or indirectly related to any physical trait (e.g., filter responses for iris images) [10]. The extracted set is commonly referred to as the template and is used as an input for matching and filtering (classification/indexing). Ideally, the extracted features are consistent for the same subject (small intra-class variation) and are distinct between different subjects (small inter-class similarity). In practice, however, factors such as poor image quality and distortion can greatly affect the accuracy of feature extraction.

Feature extraction can be related to dimensionality reduction, where the raw input signal is often in high dimension, containing redundant and irrelevant information [2]. Feature extraction transforms the original data space into a lower dimension by retaining the most discriminatory information possible. In fact, standard dimensionality algorithms (e.g., PCA) are commonly employed to extract features for face images. Regardless of the trait, the feature extraction algorithm greatly controls the performance of matching [10]. If feature extraction can separate the subjects in the feature space, simple matching algorithms can be employed. If feature extraction performs poorly, it may not be possible to design a matching algorithm that will provide sufficient accuracy.

For some applications, especially those where multiple systems need to work together, algorithms need to be interoperable. That is, in particular, the extracted features, or templates are encoded in such a way that they can be used by any matching system that follow the same encoding standard. This is crucial for large-scale applications, such as biometric passport, especially when the template storage space is small. Once again, standardization bodies play an important role in defining common formats to store the biometric templates. The Minutiae Interoperability Exchange Test (MINEX) [11], conducted by NIST, quantified the impact on system performance to use fingerprint minutiae standards in comparison to proprietary formats.

Matching

A matching algorithm compares the features extracted from the query with the stored templates in the database to produce scores that represent the (dis)similarity between the input and template. A matching algorithm must cope with variations of the extracted features [12]. These variations may be the result of modification (e.g., scar, aging, disease), occlusion (e.g., beard, glasses), presentation (pose, displacement, nonlinear distortion), and noise (lighting, motion blur) of the biometric trait. Variations resulting from the presentation of the biometric are typically handled through the use of invariant features or by trying to "align" the two templates. A common approach to alleviate some variations is to introduce certain flexibility (or tolerance) in the matching of individual features (local matching) and obtain an accumulated probability value (global matching) for computing the final match score. In many cases, this approach is shown to exhibit some complementary nature, increasing robustness to errors while preserving high accuracy. Integration (fusion) of various feature representations in a matching algorithm or combining different matching algorithms seems to be the most promising way to significantly improve the matching accuracy.

In the final stage, matching must provide a decision, either in the form of validating a claimed identity or providing a ranking of the enrolled templates to perform identification. The biometric matching algorithms range from simple nearest neighbor algorithms, to sophisticated methods such as support vector machines. Thresholding techniques are used to decide if the distance of the claimed identity (in verification) or first rank (in identification) is sufficient for authentication.

In large systems, such as countrywide ID or law enforcement systems, when throughput is high or when matching decision has to be determined online in real time (e.g., border crossing), the time of an individual match must be very small. This imposes strong constraints on the design of the matching algorithm. In order to achieve both high accuracy and speed, ► multistage matching techniques are often used. Furthermore, biometric algorithms can often be implemented in a parallel architecture, and the processing of matching can be distributed over many CPUs.

Filtering (Classification/Indexing)

With the rapid proliferation of large-scale databases, one to one matching of the query with each template in the database would be computationally expensive. A filtering process is, hence, usually employed to reduce the number of candidate hypotheses for matching operation. Filtering can be achieved by two different approaches: classification and indexing [10].

Classification algorithms, or classifiers, partition a database into a discrete set of classes. These classes can be explicitly defined based on the global features of the biometric data, e.g., "Henry classes" for fingerprints [13]; or implicitly derived based on data statistics [10]. General biometric classification algorithms can be divided into rule-based, syntactic-, structural-based, statistical- and Neural Network-based and multi-classifier methods. Sometimes, a single-level classification is not efficient enough as data may be unevenly distributed among these classes. For example, more than 90% of fingerprints belong to only three classes (left loop, right loop, and whorls). To continue narrowing down the search, some classes can be further divided into more specific categories, also known as sub-classification. Once templates in a database are classified, matching time can be greatly reduced by comparing the query only with templates belonging to the same class assigned to the query.

Indexing algorithms [10], on the other hand, provide a continuous ordering of the database. This process is also often referred as continuous classification, where biometric data are no longer partitioned into disjoint classes, but associated with numerical vector representations of its main features. This can also be regarded as an extremely fast matching process, where feature vectors can be created through a similarity-preserving transformation and the matching is performed by comparing the query only with those in the database whose vector representation are close to that of the query in the transformed space.

Because they can be extremely fast, filtering techniques are often used as a first stage in multistage matching. Indexing is often preferred over classification, since it enables to avoid classifying ambiguous data (e.g., by adjusting the size of the neighborhood considered for matching) and can be designed to be virtually error free.

Fusion

Biometric systems can be designed to recognize a person based on information acquired from multiple biometric sources. Such systems, also known as multi-biometric systems, offer substantial improvement with regard to enrollment and matching accuracy over traditional (uni) biometric systems [12, 14]. The algorithm that combines the multiple sources of information in a multibiometric system is called fusion.

Biometric fusion can be performed at four different levels of information, namely, sensor, feature, match score, and decision levels [12, 14]. Fusion algorithms can be used to integrate primary biometric traits (e.g., fingerprint and face) with soft biometric attributes (e.g., gender, height and eye color). Besides improving recognition accuracy, information fusion also increases population coverage (by avoiding "failure to enroll" and deters spoof attacks in biometric systems [14].

Related Entries

► Biometric Sample Acquisition
► Biometric, Overview

References

1. Jain, A., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 4–20 (2004)
2. Duda, R., Hart, P., Stork, D.: Pattern classification 2nd edn. Wiley, New York (2000)
3. Brislawn, C.: The FBI fingerprint image compression specification. In: Topiwala, P. (ed.) Wavelet Image and Video Compression, pp. 271–288. Kluwer (1998)

4. Brislawn, C., Quirk, M.: Image compression with the JPEG-2000 standard. In: Driggers, R. (ed.) Encyclopedia of Optical Engineering, pp. 780–785. Marcel Dekker (2003)
5. INCITS biometric sample quality standard draft. M1/06-0948 (2006) http://www.incits.org/tc_home/m1htm/2006docs/m1060948.pdf
6. NIST Biometric Quality Workshop, 2006, http://www.itl.nist.gov/iad/894.03/quality/workshop07/index.html (2007)
7. Tabassi, E., Wilson, C., Watson, C.: Fingerprint image quality. NIST research report NISTIR7151 (2004)
8. ISO/IEC Biometric Sample Quality Standard. ISO/IEC 29794 http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43583
9. Hong, L., Wan, Y., Jain, A.: Fingerprint image enhancement: Algorithms and performance evaluation. IEEE Trans. Pattern Anal. Mach. Intell. **20**(8), 777–789 (1998)
10. Maltoni, D., Maio, D., Jain, A., Prabhakar, S.: Handbook of fingerprint recognition. Springer, New York (2003)
11. Grother, P., McCabe, M., Watson, C., Indovina, M., Salamon, W., Flanagan, P., Tabassi, E., Newton, E., Wilson, C.: MINEX: Performance and Interoperability of the INCITS 378 Fingerprint Template. NIST MINEX Evaluation Report (2006)
12. Jain, A., Flynn, P., Ross, A.: Handbook of biometrics. Springer, New York (2008)
13. Henry, E.: Classification and uses of finger prints. Routledge, London (1900)
14. Ross, A., Nandakumar, K., Jain, A.: Handbook of multibiometrics. Springer, New York (2006)

Biometric and User Data, Binding of

Peng Li, Jie Tian, Xin Yang, Sujing Zhou
Institute of Automation, Chinese Academy of Sciences, Beijing, People's Republic of China

Synonyms

Key binding; Secure biometrics; Template protection

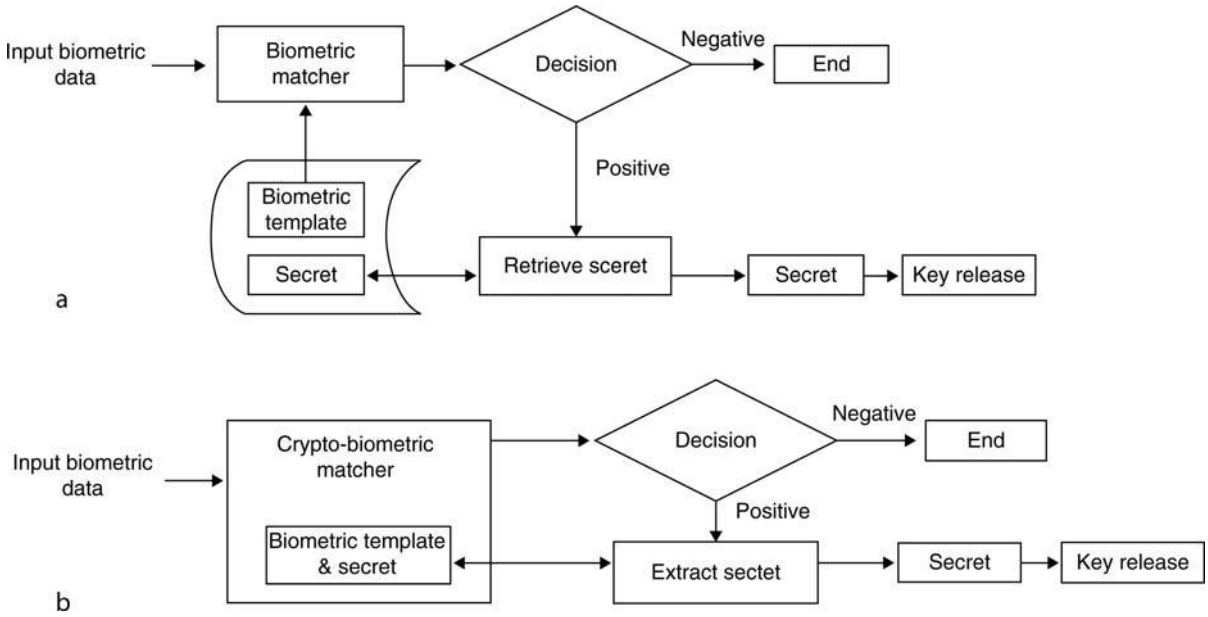
Definition

"User data" stands for the private information of the biometric system user, for example the identity number, e-mail address or any other significant or insignificant character string, which can be transformed into ASCII code in computer systems. Binding of biometric and user data is a method which aims to solve the issues of security and privacy involved with biometric system. As an important method of Biometric Encryption, binding of biometric and user data has two main functions, one is protecting the biometric template from attacks, where cancelable biometric templates can be generated; and the other is embedding user data into the biometric template, where user data will be reproduced if and only if biometric matching succeeds.

Introduction

As an identity authentication method, biometrics bases recognition on an intrinsic aspect of the human being and the use of biometrics requires the person to be authenticated as physically present at the point of the authentication [1]. With more and more application examples, biometrics recognition system exposes some intrinsic defects; the most serious is the security and privacy issue involved with raw biometric data [2]. Biometric data is difficult to cancel in case it is lost or obtained by an attacker. The lost biometric may be used for cross-matching between different databases and can bring disastrous results to user data. Because of this kind of hidden danger, people resort to a more secure biometric system, called Biometric Encryption or Biometric Cryptosystems [3]. Among the various methods of Biometric Encryption, binding of biometric and user data is the most practical and promising one, which is named Key Binding Method. This is different from the other key-related method: Key Release (Fig. 1).

The commonly collected user data includes name, any form of ID number, age, gender, and e-mail address, etc. The user data which is bound with biometric in the algorithm layer, say e-mail address or social security ID, should be protected from being stolen by the attackers, while the nonsensitive data can be open. In the enrollment stage of the Biometric Encryption system, the biometric feature extraction procedure is the same as in the traditional system. After the feature is obtained, it will be bound with user data (e.g., identity number, password, etc) in some way, thus yielding a cancelable biometric template, which will be stored as a private template and used to match the query samples. In the matching stage, the user provides his/her biometric and the user-specific data to the biometric system. Then, the same feature extraction and binding procedure will be conducted inside the system. The two private templates are compared in the traditional



Biometric and User Data, Binding of. Figure 1 (**a**) Key release (**b**) Key binding (Reprinted with permission from Jain et al.: Biometrics: a tool for information security. IEEE Trans. Inf. Forensics Security. **1**(2), 125–143 (2006) ©2006 IEEE).

manner in which the biometric system works and a matching score or YES/NO decision is given. In some algorithms, famous fuzzy vault algorithm for instance, if matching succeeds, the user-specific key is reproduced and released. The key can be used in different conventional cryptographic circumstances.

Challenges

The difficulty of binding biometric and user data lies mostly in how the fuzziness of biometrics and the exactitude of user data (key) are bridged.

Fuzziness of Biometrics

Unlike the password-based identity authentication system, biometric signals and their representations (e.g., fingerprint image and its computer representation) of a person vary dramatically depending on the acquisition method, acquisition environment, and user's interaction with the acquisition device [2].

Acquisition condition variance: The signal captured by a sensor varies with the identifier as well as the acquisition equipment. For example, fingerprint images are usually captured with contacting sensors, e.g., capacitive sensor, inductive sensor, and optical sensor. The mechanism of imaging fingerprint is mapping a three-dimensional object to a two-dimensional plane. Since the finger tip is nonrigid and the mapping procedure cannot be controlled precisely, the captured fingerprint images change in minute details from time to time, but are still within a certain metric distance of intra-class difference. When the sensor's surface is not large enough or the user provides only part of the finger to the sensor, the acquired image area does not cover the whole finger. Different fingerprint images from the same finger may include different parts of the finger. In addition, translation and rotation are very common in different samples from the same finger. Another good example to show the acquisition condition variance is facial image acquisition. Illumination change influences the captured facial image in real circumstances. Moreover, the greatest variance is in the facial expression, including the kinds of modalities used to express different emotions. Almost all kinds of biometric modalities have to bear this variance.

Circumstances and time variance: Change in outer circumstances may also cause the captured biometric signal to vary more or less. While taking the fingerprint, for example, the environmental temperature and humidity may render the finger too dry or too

damp to be captured. Low-quality fingerprint images are very common in real application systems and enhancing (i.e., preprocessing) them is a challenging research direction in the traditional fingerprint recognition field. Generally, the fingerprint does not change with time because the skin on the finger tip may not change much with age. But many modalities cannot resist the temporal change, e.g., face, gait, palm, voice, and so on. In particular, the face varies greatly with age; facial images captured from the same person at different ages differ vastly. How one estimates the aging model of a person also makes an important research issue in the face recognition field. In addition, there are other factors which can influence the captured biometric signal for some specific modality.

Feature extraction variance: Almost all the feature extraction algorithms are based on signal processing or image processing methods. They are not exact when processing different biometric samples. Noise is often introduced in the extraction procedure, especially of the low-quality samples.

All the above factors can make the samples from the same subject seem different and the ones from different subjects quite similar. Large intra-class differences and small inter-class differences will be the result due to these reasons. However, a cryptosystem requires exact computing and operation. A tiny change in input may cause an enormous difference in output, for example, for the hash function. So bridging the fuzziness of biometrics and the exactness of cryptography becomes the greatest challengein the binding of biometric and user data.

Encrypted Template Alignment

The second challenging problem is how to align the encrypted biometric templates. One of the purposes of binding biometric and user data is to conceal raw biometric data. Thus original features cannot be used for alignment after binding to prevent the original template from being stolen. Nevertheless, the alignment stage has to be conducted to locate the various biometric samples in the same metric space and to ensure the authentication accuracy. So the feature used in the alignment stage must satisfy two conditions: (1) it will not reveal original biometric data and (2) it must assure alignment accuracy to some extent. The concept of Helper Data satisfying these conditions was proposed [4]. Taking the case of the fingerprint as an example, the points with maximum local curvature around the core are detected and used for alignment without leaking the minutiae information. Theoretically, the system security can be estimated according to information theory from the information published by Helper Data [5].

Theory and Practice

The theories of Secure Sketch [5] and Fuzzy Extractor [5] lay the foundation for the binding of biometric and user data and give some significant theoretic results from the point of view of information theory. In the various binding methods of biometric and user data, Bioscrypt [6], Biohashing [7], Fuzzy Commitment [8], and Fuzzy Vault [9] are the most representative to address the problem of security and privacy. These algorithms will be described in detail in the next section.

Fuzzy Commitment Scheme

Fuzzy Commitment scheme [8] is one of the earliest methods of binding biometric and user data. It is actually an ordinary commitment scheme (a primitive in cryptography) taking biometric templates as private keys, and employing error correcting codes to tackle the fuzziness problem of biometric templates.

As an ordinary cryptographic commitment, the fuzzy commitment scheme has two procedures: committing and decommitting. To commit a bit string x, first generate a codeword c from x according to a prespecified error correcting code, then apply some cryptographic hash function (or one-way function) to c, the ultimate commitment is (h(c), w + c), where w is a biometric template related string with the same length of c. To decommit a commitment, the user has to provide a biometric template related string w' which is close to that in the committing procedure; the verifier uses it to decode the correct codeword c, then checks whether the hash value of c equals the stored hash value in the commitment, and accepts the commitment if they are equal, rejects otherwise. The fuzzy commitment scheme is essentially a Secure Sketch as observed by Dodis et al. [5].

Secure Sketch

A Secure Sketch is a primitive component proposed by Dodis et al. [5] to extract helper data from the input

biometric sample and to reconstruct the original sample according to the helper data without storing the raw biometric template. A Secure Sketch consists of two procedures. The first procedure outputs a bit string (called helper data) from the enrolled biometric template and stores the bit string while discarding the enrolled biometric template. In the second procedure, the query sample is inputted. The biometric template could be reconstructed according to the query and the helper data if the distance between the query and the template is less than a specified threshold in terms of some metric space.

The security of a Secure Sketch is estimated as the loss of the min-entropy of the enrolled biometric template between the sketch values before and after thebit string is provided; the less the loss the better. In case the distance between the two biometric templates is measured by the number of positions in which the two binary represented biometric templates differ, e.g., in Hamming metric space, two basic constructions based on error correcting codes are known: code-offset construction and syndrome construction. In case the distance between the two biometric templates is measured by the number of elements that occur only in one of the two duplicate-free set represented biometric templates, e.g., in set difference metric space, the construction is called a PinSketch. In case the distance between two biometric templates is measured by the smallest number of character insertions and deletions required to change one biometric template into another one, e.g., in edit metric space, the metric space is first transformed into another metric space that is easy to handle by embedding injections with some distortion that is tolerable, and then treated as in the transformed metric space.

A Secure Sketch can be used to construct a fuzzy extractor. Fuzzy vault and fuzzy commitment in this context are essentially Secure Sketches in Hamming metric space and set difference metric space, respectively.

Fuzzy Extractor

A fuzzy extractor is a primitive component proposed by Dodis et al. [5] to obtain a unique bit string extracted from the biometric template provided in enrollment whenever the query biometric template is close enough to the enrolled biometric template. The random bit string can be further used as a private key of the user.

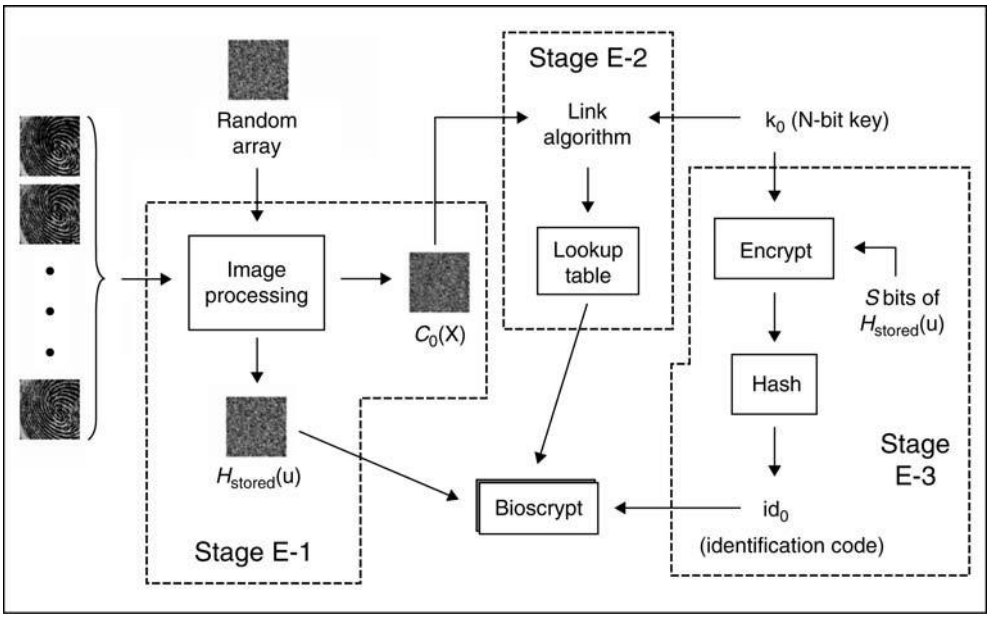
A fuzzy extractor consists of two procedures. The first procedure outputs a bit string and a helper data from the enrolled biometric template, stores the helper data while it discards the bit string and the enrolled biometric template. The second procedure outputs the bit string from the first procedure if the distance between the query biometric template and the enrolled biometric template is less than a specified parameter, given a query biometric template and helper data from the first procedure.

The security of a fuzzy extractor is estimated as the statistics distance between the bit string and a uniform random distribution when the helper data is provided; the closer the better. A fuzzy extractor can be constructed easily from any Secure Sketch. A fuzzy extractor itself is also an important primitive component in biometric based cryptosystems. Fuzzy extractors with robustness [5] are considered to protect against a kind of active attack, i.e., an adversary might intercept and change the helper data in a way to obtain biometric template-related private information of the user who blindly applied his biometric template on the fraud helper data. Fuzzy extractor with reusability [5] is also considered to secure against a kind of active attack, i.e., a collusion attack from multiple application servers to which a user is enrolled by the same fuzzy extractor scheme, each server obtaining a different helper data and by collusion there exists the risk of exposure of private user data, e.g., biometric template.

Bioscrypt

Bioscrypt [6], a method of binding biometric and user data, is the first practical Biometric Encryption algorithm to the authors' knowledge. The binding is based on performing a Fourier Transform of a fingerprint.

In the enrollment stage (Fig. 2), several fingerprint images, denoted by $f(x)$ are inputted and Fourier Transformation and other operations are performed to result in $H(u)$. $H(u)$ composes two components: magnitude $|H(u)|$ and phase $e^{i\varphi_H(u)}$. The magnitude component $|H(u)|$ is discarded and the phase $e^{i\varphi_H(u)}$ is preserved. A random array is generated according to RNG (Random Number Generator), denoted by R. The phase components of R, denoted by $e^{i\varphi_R(u)}$, are used to multiply with $e^{i\varphi_H(u)}$ and results stored in $H_{stored}(u)$. In addition, $c_0(x)$ is produced from the



Biometric and User Data, Binding of. **Figure 2** Overview of the enrollment process for Biometric Encryption. (Reprinted with permission from [5]).

Fourier Transformation of the number of fingerprints and stored into a lookup table together with an N-bit key k_0, where k_0 and $c_0(x)$ are linked with a link algorithm. On the other hand, k_0 is used to encrypt S bits of $H_{stored}(u)$ and then the result will be hashed to obtain an identification code id_0. After the above procedure, $H_{stored}(u)$, the Lookup table, and the identification code id_0 are stored together in a template (called Bioscrypt by the authors).

In the verification stage (Fig. 3), after inputting the query fingerprint sample and the Fourier Transformation operation, the identification code $c_1(x)$ is computed according to the $H_{stored}(u)$ in the Bioscrypt. Through the link algorithm, a key k_1 is released from the Lookup table in the Bioscrypt. id_0 is released synchronously to be used for comparing in the next step. S bits of $H_{stored}(u)$ is encrypted with k_1 and the result is hashed to result in id_1. id_1 is compared to id_0 and if they are identical the identification succeeds, otherwise fails.

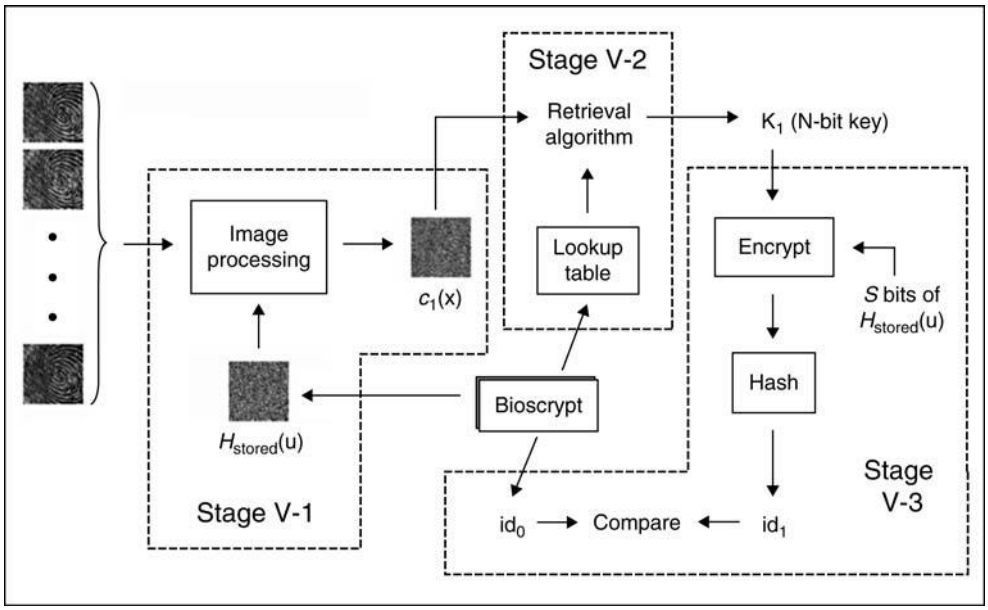
Biohashing

Biohashing [7] is also a typical Biometric Encryption algorithm binding biometric and user data. In the beginning, it uses the fingerprint, followed by face-hashing [10], palmhashing [11], and so on.

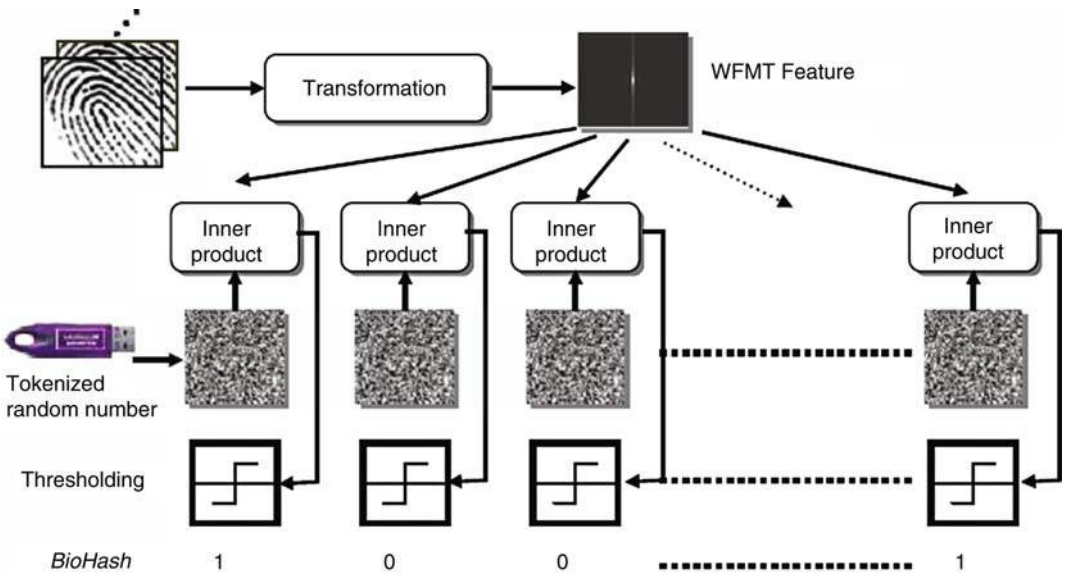
Toeh et al. [7] proposed the two-factor identity authentication method combining fingerprint and tokenized random number (i.e., user data). The Wavelet Fourier Mellin Transformation (WFMT) feature of fingerprint is employed (Fig. 4) and iterative inner product operations are performed on WFMT and the user-specific pseudo-random number stored in the user's token (Fig. 5). Quantization is then conducted on the inner product value according to the preset threshold. Thus, from a fingerprint image a bit-string can be obtained, which is used for matching in terms of Hamming distance.

However, the authentication performance of biohashing will decrease greatly if the token (i.e., the user data) is stolen by the attacker, which is called the token-stolen scenario. Related experiments have confirmed this point. That is to say, tokenized random number plays a more important role than the biometric itself in the biohashing algorithm.

Some subsequent work has focused on improving the performance in the token-stolen scenario, e.g., Lumini and Nanni's work [12], which are briefly described below. They improved the performance by dramatically increasing the length of the biohashing



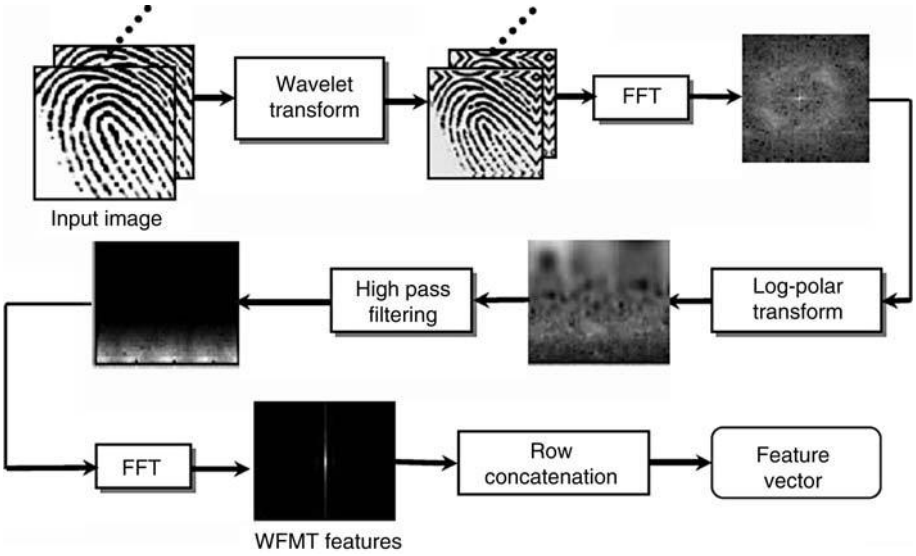
Biometric and User Data, Binding of. **Figure 3** Overview of the verification process for Biometric Encryption. (Reprinted with permission from [5]).



Biometric and User Data, Binding of. **Figure 4** The flowchart of WFMT generation. (Reprinted with permission from [6]).

output. The following are the specific solutions leading to the reported improvement:

1. *Normalization*: Normalizing the biometric vectors by their module before applying the BioHashing procedure, such that the scalar product $<x|or_i>$ is within the range $[-1, 1]$
2. τ *Variation*: Instead of using a fixed value for τ, using several values for $\tau$and combining with the "SUM rule" the scores obtained by varying τ between τ_{max} and τ_{min}, with p steps of $\tau_{step} = (\tau_{max}\text{-}\tau_{min})/p$
3. *Spaces augmentation*: Since the dimension of the projection space m cannot be increased at will, using more projection spaces to generate more



Biometric and User Data, Binding of. **Figure 5** The flowchart of iterative inner product operation. (Reprinted with permission from [6]).

BioHash codes per user. Let k be the selected number of projection spaces to be used; the Biohashing method is iterated k times on the same biometric vector in order to obtain k bit vectors b_i, $i = 1, 2, \ldots, k$. Then the verification is carried out by combining the classification scores obtained by each bit vector (BioHash code). The random generation can be performed in an iterative manner, thus requiring a single Hash key K: in such a way that the random generator is not reinitialized by a new key until the complete generation of the k bases is not performed

4. *Features permutation*: Another way to generate more BioHash codes, without creating more projection spaces, is to use several permutation methods of the feature coefficients in x during the projection calculation: using q permutations of x obtained by round-shifting the coefficients of a fixed amount thus obtaining q bit vectors. As above the verification is carried out by combining the classification scores obtained by each bit vector

Fuzzy Vault

The Fuzzy Vault algorithm [9] is a practical method of binding biometric and user's private key. It consists of the following two steps:

1. A user Alice places a secret (K) in the vault, and locks it with an unordered set A
2. Another user Bob tries to access the secret (K) with another unordered set B (i.e., unlock the vault)

Bob can access the secret (K) if and only if the two unordered sets B and A overlap substantially.

Specifically, the Fuzzy Vault can be depicted as follows:

1. *Encoding the Vault*: A user Alice selects a polynomial p of variable x encoding K, then computes the project $p(A)$ of the unordered set A on the polynomial p, thus $(A, p(A))$ can construct a finite point set. Some chaff points are then randomly generated to form R with the point set $(A, p(A))$; R is the so-called *Vault*. The chaff point set is vital to hide the secret K, and the point numbers in it are more than the real point set
2. *Decoding the Vault*: Another user Bob tries to access the secret (K) with another unordered set B. If the elements in B and the ones in A overlap substantially, then many points in B will lie in the polynomial p. So Bob can use correction code technology to reconstruct p, and consequently access the secret K. However, if a large proportion of points in B and A do not overlap, due to the difficulty of reconstructing the polynomial, it is almost infeasible to attain p over again.

The security of Fuzzy Vault scheme is based on the polynomial reconstruction problem. This scheme is

highly suitable for hiding biometric data, because it works with unordered sets (e.g., fingerprint minutiae), and can tolerate difference (element number or kind or both) between the two sets A and B to some extent.

The idea of "fuzzy fingerprint vault" [13] and "fuzzy vault for fingerprint" [4] are also proposed aiming to solve the problems of fingerprint template protection. Fuzzy Vault for face [14] and iris [15] have also been proposed recently.

Performance Evaluation

Performance evaluation of the binding of biometric and user data should be conducted based mainly on two aspects: accuracy and security. Accuracy reflects the effect after binding of biometric and user data as an enhanced identity authentication way, and security can provide information on the probability that the system will be attacked successfully.

1. **Accuracy:** The accuracy of biometric-like identity authentication is due to the genuine and imposter distribution of matching. The overall accuracy can be illustrated by Receiver Operation Characteristics (ROC) curve, which shows the dependence of False Reject Rate (FRR) on False Accept Rate (FAR) at all thresholds. When the parameter changes, FAR and FRR may yield the same value, which is called Equal Error Rate (EER). It is a very important indicator to evaluate the accuracy of the biometric system, as well as binding of biometric and user data.
2. **Security:** The security of the binding of biometric and user data depends on the length of user data, which is converted to binary 0/1 expression. It assumes the attacker has full knowledge about the binding method, but can only mount brute-force attack on the system. So the system security is weighed by bit length of the user data. Typically, the security of the iris binding system is 140-bit,and that of fingerprint is 128-bits. However, typical face binding algorithm holds only 58-bit security [3].

Summary

Binding of biometric and user data is a kind of technique to tackle the issues of security and privacy arising frequently in traditional biometric systems. It may decrease the accuracy performance to some extent, but generally, the security and privacy of the system are enhanced.

Related Entries

- ▶ Privacy Issues
- ▶ Security Issues, System Design

References

1. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003)
2. Uludag, U., Pankanti, S., Prabhakar, S., Jain, A.K.: Biometric cryptosystems: issues and challenges. Proc. IEEE. **92**(6), 948–960 (2004)
3. http://www.ipc.on.ca/images/Resources/up-1bio_encryp.pdf.
4. Uludag, U., Jain, A.: Securing fingerprint template: fuzzy vault with helper data. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop, New York, pp. 163 (2006)
5. Dodis, Y., Reyzin, L., Smith, A.: Fuzzy extractor. In: Tuyls, P., Skoric, B., Kevenaar, T. (eds.) Security with Noisy Data. Springer, London (2008)
6. Soutar, C., Roberge, D., Stojanov, S.A., Gilroy, R., Kumar, B.V.K.V.: Biometric encryption. In: Nichols, R.K. (ed.) Proceedings of ICSA Guide to Cryptography. McGraw-Hill, New York (1999)
7. Teoh, A.B.J., Ngo, D.C.L., Goh, A.: Biohashing: two factor authentication featuring fingerprint data and tokenised random number. Pattern Recognit. **37**(11), 2245–2255 (2004)
8. Juels, A., Wattenberg, M.: A fuzzy commitment scheme. In: Proceedings of ACM Conference on Computer and Communications Security (CCS), Singapore, pp. 28–36 (1999)
9. Juels, A., Sudan, M.: A fuzzy vault scheme. In: Proceedings of IEEE International Symposium on Information Theory, Lausanne, Switzerland, pp. 408 (2002)
10. Ngo, D.C.L., Toeh, A.B.J., Goh, A.: Eigenface-based face hashing. In: Proceedings of International Conference on Biometric Authentication, Hong Kong, China, pp. 195–199 (2004)
11. Connie, T., Teoh, A., Goh, M., Ngo, D.: PalmHashing: a novel approach for cancelable biometrics. Inf. Process. Lett. **93**(1), 1–5 (2005)
12. Lumini, A., Nanni, L.: An improved biohashing for human authentication. Pattern Recognit. **40**(3), 1057–1065 (2007)
13. Clancy, T.C., Kiyavash, N., Lin, D.J.: Secure smartcard-based fingerprint authentication. In: Proceedings of ACM SIGMM 2003 Multimedia, Biometrics Methods and Applications Workshop, Berkley, CA, pp. 45–52 (2003)
14. Nyang, D., Lee, K.: Fuzzy face vault: how to implement fuzzy vault with weighted features. In: Proceedings, Part I of Fourth International Conference on Universal Access in

Human-Computer Interaction, UAHCI 2007, Held as Part of HCI International 2007, Beijing, China, pp. 491–496. Springer, Heidelberg (2007)
15. Lee, Y.J., Bae, K., Lee, S.J., Park, K.R., Kim, J.: Biometric key binding: fuzzy vault based on iris images. In: Proceedings of the Second International Conference on Biometrics, Soul, Korea, pp. 800–808. Springer, Heidelberg (2007)

Biometric Applications, Overview

David Day
International Biometric Group, New York, NY, USA

Synonym

Biometrics

Definition

A biometric application is the sum of the functionality, utilization, and role of a biometric technology in operation. Biometric technologies such as fingerprint, face recognition, and iris recognition are utilized in a range of applications that vary in terms of performance requirements, operational environment, and privacy impact. Biometric technology selection – which modality to utilize and what hardware and software to deploy – is typically driven in large part by the application. Biometric applications can be generalized into four categories. The first application category is controlling access to data, such as logging into a device, PC, or network. The second application category is controlling access to tangible materials or areas, such as physical access control. The third application category is to validate a claimed identity against an existing credential, such as in a border control environment. The fourth application is to register or identify individuals whose identities need to be established biometrically, most often using centralized or distributed databases. Beyond this high-level decomposition, an application taxonomy can be defined that spans 12 distinct biometric applications. This taxonomy takes into account factors such as the user's motivation and incentive, the location of biometric data storage and matching, the nature of the data or materials that the biometric is protecting, and the role of non-biometric authentication and identification techniques.

Introduction

The need for secure, reliable identity validation and confirmation has driven the adoption of biometric technologies in a diverse range of applications. Biometric applications can be generalized into four categories. The first application category is controlling access to data, such as logging into a device, PC, or network. The second application category is controlling access to materials or areas, such as physical access control. The third application category is to validate a claimed identity against an existing credential, such as in a border control environment. The fourth is to register or identify individuals whose identities need to be established by biometric means, most often using centralized or distributed databases. Law enforcement and military uses of biometrics are primary examples of this fourth application category.

Though the four generalized functionalities provide an overview of how biometrics can be applied, a more detailed taxonomy is required to capture the full breadth of biometric application. The large majority of biometric utilization and ▶ deployment can e grouped into one of twelve applications:

- *Law Enforcement (forensics): Biometric technologies have long been utilized as a secure means to identify alleged criminals. In this particular application, an individual's fingerprints are used to determine or confirm an identity against a central record store. The FBI currently holds one of the largest biometric databases, comprised of tens of millions of civil and criminal fingerprint records.*
- *Background Checks: Biometric technologies are used to execute background checks as a condition of employment for many government and commercial professions. While background checks may be executed against the same databases used in criminal searches, the applications differ in that background check or "civil" records are typically not retained – they are discarded after the result is returned to the querying agency.*
- *Surveillance: Biometric technologies are deployed locate, track, and identify persons in a field of view (i.e., in a given space or area). Historically,*

surveillance applications required laborious and monotonous monitoring of cameras. Biometrics automates the process through the utilization of face recognition technology; biometric surveillance systems can be configured to alert officials to the presence of individuals of interest.4

- *Border Control: The ever-increasing volume of international travel necessitates implementation of technologies that can automate, streamline, and expedite border crossing. Driven by international standards for biometric-enabled passports, as well as ad hoc regional efforts, countries utilize fingerprints, iris, and face recognition technologies in border control applications ranging from localized to nationwide. Deployed properly, biometrics can ensure that screening resources are routed toward travelers whose risk profile is unknown.*
- *Fraud Reduction: Biometric technology can be deployed in public-sector applications to prevent individuals from claiming benefits under multiple identities. Government agencies have utilized iris and fingerprint recognition as a means to deter "double dipping" at the state and federal levels.*
- *Trusted Traveler: This application enables users to traverse security checkpoints with reduced likelihood of rigorous security inspections. Iris recognition and fingerprint are the leading technologies in this high-profile biometric application.*
- *Physical Access Control: Physical access control is use of biometrics to identify or verify the identity of individuals before permitting access to an area. Companies and government agencies deploy technologies such as fingerprint, hand geometry, and iris recognition to control key entry and exit points.*
- *Time and Attendance: Biometrics can serve as a commercial application to assist in employee management. In this particular application, devices are used to track employee attendance. Hundreds of commercial deployments utilize hand geometry and fingerprint recognition to ensure the integrity of work hours and payroll.*
- *Consumer Recognition: This application refers to the confirmation of one's identity in order to execute a commercial transaction. Conventional authentication methods have utilized keycards, PIN numbers and signatures to ensure the validity of a given transaction. Biometrics can reduce reliance on tokens and passwords and can provide consumers with a sense of assurance that their transactions are secure. Fingerprint recognition is a common technology deployed in this application.*
- *Remote Authentication: Biometrics provide a secure method of authentication for remote access to important information by allowing mobile device users to be accurately identified.1 Previous deployments have utilized fingerprint and voice recognition.*
- *Asset Protection: This application describes the need to protect digital information and other sensitive materials from unauthorized users. One common application is the use of fingerprint recognition on safes to protect sensitive documents. Biometrics also serves to compliment already in-place security methods such as passwords and user identification on computer workstations.*
- *Logical Access Control: Biometrics is used to control access to systems and/or devices based on physical characteristics. It is commonly used to control access to centralized databases, healthcare information, or financial records. Many deployments have utilized fingerprint recognition due to its proven reliability, ease-of-use, and accuracy.*

As seen by the aforementioned application descriptions, biometric technology is typically used in applications where it can improve security, increase efficiency, or enhance convenience. Additionally, biometrics allow users to forego the responsibility of creating passwords and carrying keycards while maintaining a level of security that meets, and in some cases surpasses, that of conventional authentication methods.

Discussion

Each application utilizes biometrics as a solution to an identified authentication problem. There exist, however, key differentiating factors that help to distinguish one application from another. Some of these distinctions include the environment in which biometrics has been implemented, the purpose that biometrics is intended to serve, and the methods in which biometrics is utilized to serve its purpose.

The application of biometrics in law enforcement has utilized fingerprint recognition as a reliable means to identify criminals. Biometrics enable officials to conduct automated searches, compare biometric information of suspects against local, state, and national databases, and process mug shot-database

comparisons.[3] A typical deployment would utilize a live-scan system, AFIS technology including matching hardware and software, and face recognition software. Though such a scenario is common for law enforcement related applications, recent trends have begun to push for mobile biometric devices in order to identify individuals in the field without the need to retain suspects for extended periods of time. Law enforcement applications of biometrics are unique in that they implement widely-adopted standards for imaging, data transmission, and file formats. These standards allow jurisdictions to share fingerprint and face data in an interoperable fashion, even when biometric hardware and software are sourced from different suppliers. Increasingly, law enforcement biometric systems are deployed to search suspected terrorist data as well as data collected in military applications.

Background checks utilize biometric systems to determine the identity of an individual and to retrieve his or her historical records. Biometric background check systems collect high-quality fingerprints for submission to state or federal systems that determine whether a given set of fingerprints is linked to criminal or other derogatory records. For example, some government agencies require individuals to submit biometric data for employment purposes. Fingerprint recognition technology is primarily used due to the extensive collection of fingerprint images currently held by government officials.

Surveillance applications utilize biometric technology, primarily face recognition, to locate and identify individuals without their awareness. Such applications are designed to collect biometric data without an explicit, direct presentation. By contrast, fingerprint and vein recognition technologies require individuals to voluntarily submit biometric measurement to the device. Surveillance application can, however, measure one's biometrics from a distance. A typical deployment would be to implement biometrics into already-existing security cameras or to install customized cameras whose resolution and performance characteristics are sufficient for acquisition of enrollable face images. In the future, gait recognition is envisioned as a surveillance technology capable of operating at greater distances than face recognition.[2] The technology could then notify officials to the presence of specific individuals in highly trafficked areas such as airport terminals. One challenge facing biometric surveillance is individual movement. Previously deployed systems have shown that quick and sudden actions can cause recognition performance to decrease. Some implementers have attempted to overcome this challenge by installing cameras in locations in which movement is limited such as entrances and staircases.

Border control focuses upon the management of international borders at targeted locations. At busy points of entry, it can be a difficult process to accurately and efficiently identify individuals. A common solution is to compliment conventional security protocols, such as identification cards, with biometric security methods such as fingerprint devices. This allows for 1:1 biometric matches that can reduce the time required to confirm the user's identity. There are some complications, however, when implementing biometrics into border management. One challenge typically faced is the assurance of cross-jurisdictional interoperability. It can prove to be difficult to have bordering nations to agree upon a single standard.

Biometric technology can provide a considerable financial benefit to both the government and general public. Biometric systems are deployed in public services applications for fraud reduction, detecting and deterring the use of multiple identities to receive entitlements such as welfare payments. If a previously enrolled individual attempts to claim another identity, the biometric system recognizes this and officials are alerted. Past deployments have utilized stationary fingerprint or iris recognition systems that have been installed within government facilities.

The trusted traveler application enables frequent travelers to bypass extensive and time consuming security check points after their initial enrollment. At enrollment, passengers submit their identification information and biometric data, which is then used to conduct a background check. Once the individual has been cleared as non-threatening and their identity is verified, the agency can then distribute a specialized traveler's smart card that contains the traveler's information and biometric data.[3] With this smart card, the traveler can utilize specialized security checkpoints to gain access to airport terminals quickly and conveniently. Terminals install automated systems that determine whether to deny or grant access to the traveler based on their biometric information. Typical trusted traveler systems utilize gated entry points to prevent forced entry, smart cards that store biometric templates, and face, fingerprint, or iris recognition technology to verify the individual's identity. The

commercial benefits of trusted traveler programs accrue when a critical mass of registrants is reached, as well as when additional programs are incorporated into the "trusted" framework.

Biometric physical access control deployments are most often implemented to control employee access to secure or protected areas. Typically, the biometric reader is installed as a stationary system in which the user must verify his or her identity against a card-based, reader-based, or centralized template. Physical access control is one of the most well-established biometric applications, with hundreds of devices on the market ranging from inexpensive, standalone fingerprint readers to highly automated iris recognition devices. Fingerprint, face recognition, hand geometry, and vein recognition are also commonly deployed for physical access control.

Aside from maintaining a high level of security, biometric applications can help to serve the commercial sector for financial benefits. Biometrics used for time and attendance confirm the presence of an individual at a specific time, date, and location. Because of the potential difficulty of tracking the hours of thousands of employees at larger facilities, time and attendance applications allow management to automatically eliminate the possibility of "buddy-punching", tardiness, or absence without their knowledge. Automating this process can also lead to time savings with payroll management. Hand geometry recognition and fingerprint are the most-frequently deployed biometric technologies in this application. Deployers often need to overcome the learning curve associated with device acclimation and the challenge of end-user acceptance.

Biometrics are deployed in financial sector applications to provide convenience and security for the consumer. Numerous banks have deployed fingerprint and vein recognition technology at ATMs as a method to enhance identification and security. The use of biometric technology bypasses the need for users to carry identification cards and to remember lengthy PIN numbers. Another possible application of biometric technology within the financial sector is to use biometrics in customer service call centers. This specific example utilizes voice recognition technology to bypass the need for customers to provide their identification details and verify their information. Instead, voice recognition technology automates the customer authentication process, and allows representatives to immediately aid the consumer, saving time and increasing productivity.

Remote authentication utilizes biometrics to verify individuals in different locations, and allows for unsupervised secure authentication. Web-based financial transactions without biometrics typically consist of an extended identification number, PIN number, and/or user information to verify an individual as authorized. This single factor authentication, however, can be easily replicated. Biometric technologies such as voice recognition or mobile fingerprint recognition can provide an added layer of security to reduce customer fraud.

Biometric logical access control applications allow authorized users to gain access to systems or devices containing highly sensitive information such as healthcare information and financial records. A common approach would be to utilize inexpensive fingerprint peripherals (for workstations) or integrated fingerprint devices (for laptops). The user must provide his or her biometric information in order to gain access to sensitive device or system. It can be a challenge to deploy biometrics for logical access control because end-users may feel uncomfortable with supplying such personal information to gain access to information. It would be crucial to provide sufficient lead time for users to become accustomed wit the device and aware of what information is being recorded and not recorded.

Summary

Biometrics technologies are currently deployed in a wide range of mission-critical government and commercial applications. Due to its wide range of functionality, biometric technology can be utilized in a number of applications to provide an added-level of security and convenience beyond that of conventional security methods. Additionally, biometrics can be implemented in parallel with legacy systems to enable a gradual transition from conventional security systems to enhanced biometric security. As seen from previous deployments, some biometric modalities better serve one application than another; limiting factors include environment, size, and end-user compliance. Though each application serves its own purpose, applying biometrics achieves the overarching goal of accurately identifying or verifying an individual's identity while enhancing security, efficiency, and/or convenience.

Related Entries

- ▶ Access Control, Logical
- ▶ Access Control, Physical
- ▶ ePassport
- ▶ Law Enforcement
- ▶ Surveillance
- ▶ Time and Attendance

References

1. Lee, C., Lee, S., Kim, J., Kim S.: Preprocessing of a fingerprint image captured with a mobile camera. Biometrics Engineering Research Center, Korea (2005)
2. Lu, H., Plataniotis, K., Ventsanopoulos, A.: Uncorrelated Multilinear Discriminant Analysis with Regularization for Gait Recognition. The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, Canada (2007)
3. Arnold, M., Busch, C., Ihmor, H.: Investigating performance and impacts on fingerprint recognition systems. United States Military Academy, New York (2005)
4. Sellahewa, H., Jassim, S.: Wavelet based face verification for constrained platforms. In: Proceedings of SPIE 5779. Florida, 2005

Biometric Capture

It refers to the stage of the biometric authentication chain in which the biometric treat is transformed into an electrical signal, which is useful for further processing.

- ▶ Biometric Sensor and Device, Overview

Biometric Capture Device

Biometric capture device is a device that captures signal from a biometric characteristic and converts it to a digital form (biometric sample) suitable for storing, and automated comparison with other biometric samples.

- ▶ Fingerprint Image Quality

Biometric Characteristic

Biological and behavioral characteristic of an individual that can be detected and from which distinguishing, repeatable biometric features can be extracted for the purpose of automated recognition of individuals. Biological and behavioral characteristics are physical properties of body parts, physiological and behavioral processes created by the body and combinations of any of these. Distinguishing does not necessarily imply individualization (e.g., Galton ridge structure, face topography, facial skin texture, hand topography, finger topography, iris structure, vein structure of the hand, ridge structure of the palm, and retinal pattern).

- ▶ Multibiometrics and Data Fusion, Standardization

Biometric Cryptosystem

Biometric cryptosystems refer to systems which can be used for securing a cryptographic key using some biometric features, for generating a cryptographic key from biometric features, or to a secure biometric template. Specifically, the following operational modes can be identified. In the *key release* mode the cryptographic key is stored together with the biometric template and the other necessary information about the user. After a successful biometric matching, the key is released. In the *key binding* mode, the key is bound to the biometric template in such a way that both of them are inaccessible to an attacker and the key is released when a valid biometric is presented. It is worth pointing out that no match between the templates needs to be performed. In the *key generation* mode, the key is obtained from the biometric data and no other user intervention besides the donation of the required biometrics is needed.

- ▶ Encryption, Biometric
- ▶ Iris Template Protection
- ▶ Security Issues, System Design

Biometric Data

Biometric data, also called biometric sample or biometric record, is any data record containing a biometric sample of any modality (or multiple modalities), whether that data has been processed or not. Biometric data may be formatted (encoded) in accordance with a standard or may be vendor specific (proprietary) and may or may not be encapsulated with the metadata. Examples of biometric data include a single compressed fingerprint image, a four-finger slap image formatted as an ANSI/NIST ITL1-2000 Type-14 record, a record containing fingerprint minutiae from the right and left index fingers, a digital passport face photo, two iris images within a CBEFF structure, or an XML encoded voice sample.

► Biometric Interfaces

Biometric Data Acquisition

► Biometric Sample Acquisition

Biometric Data Block (BDB)

► Biometric Data Interchange Format, Standardization

Biometric Data Capture

► Biometric Sample Acquisition

Biometric Data Interchange Format

► Biometric Data Interchange Format, Standardization

Biometric Data Interchange Format, Standardization

Christoph Busch[1], Greg Canon[2]
[1]Fraunhofer Institute Graphische Daten Verarlaitung, Fraunhoferstrasse, Darmstadt, Germany
[2]Cross Match Technologies, RCA Blvd, FL

Synonyms

Biometric Data Interchange Format; Biometric Data Block (BDB); Biometric Reference

Definitions

Biometric Data Interchange Formats define an encoding scheme according to which biometric data is stored in a ► biometric reference. In most cases the stored data will be used for future comparisons with biometric data stemming from the same or different subject. Encoded data should not only contain a digital representation of a ► biometric characteristic (e.g., fingerprint image, face image) but also relevant metadata that impacted the capturing process (e.g., resolution of fingerprint sensor). Standardized Data Interchange Formats are a fundamental precondition to implement open systems where biometric data can be processed with components of different suppliers.

Introduction

Biometric systems are characterized by the fact that essential functional components are usually dislocated. While the enrolment may take place as part of an employment procedure in a personal office or at a help-desk, the biometric verification often takes place at different location and time. This could occur when the claimant (the data subject) approaches a certain physical access gate or requests logical access to an IT system. No matter whether the recognition system operates in verification or identification mode, it must be capable to compare the fresh biometric data captured from the subject with the stored reference data. Applications vary in the architecture, especially with respect to the storage of the biometric reference.

Some applications store the reference in a database (either centralized or decentralized), while other applications use token-based concepts like the ePassport [1] in which subjects keep control of their personal biometric data as they decide for themselves whether and when they provide the token to the controlling instance [2].

The recognition task is likely to fail if the biometric reference is not readable according to a standardized format. While closed systems that are dedicated to specific applications – say access control to a critical infrastructure – could be designed on proprietary format standards only, any open system implementation requires the use of an interoperable, open standard to allow for enrolment and recognition components to be supplied from different vendors. An operator should also be able to develop a system such that generators (and also verifiers) of biometric references can be replaced – should one supplier fail to guarantee service. Furthermore, it is desired that the same biometric reference could be used in different applications: It may serve as a trusted traveler document or as ID for eGovernment applications. Applications that may be quite different in nature will require the biometric data to be encoded in one harmonized record format. Due to the nature of the different biometric characteristics being observed, an extensive series of standards is required. Some biometric systems measure stable biological characteristics of the individual that reflect anatomical and physiological structures of the body. Examples of these types are facial or hand characteristics. Other biometric systems measure dynamic behavioral characteristics, usually by collecting measured samples over a given time span. Examples are signature/sign data that is captured with digitizing tables or advanced pen systems or voice data that is recorded in speaker recognition systems. The ISO/IEC JTC1/SC37 series of standards known as ISO/IEC IS 19794 (or the 19794-family) meets this need. This multipart standard includes currently 13 parts and covers a large variety of biometric modalities ranging from finger, face, iris, signature, hand-geometry, 3-D face, voice to DNA data.

Many applications embed the biometric data as a ▶ Biometric Data Block (BDB) in a data container such as the Common Biometric Exchange Format Framework (CBEFF) [3] that provide additional functionality such as integrity protection of the data through digital signatures or the storage of multiple recordings from various biometric characteristics in one data record. Thus the CBEFF container is also appropriate to represent data for multimodal biometric systems. The BDB is a concept described in the CBEFF standard. The CBEFF standard is a component of the SC37 layered set of data interchange and interoperability standards.

Format Structures

The prime purpose of a biometric reference is to represent a biometric characteristic. This representation must allow a good biometric performance when being compared to a fresh verification sample as well as allowing a compact coding as the storage capacity for some applications (e.g., the RFID token with 72 Kbytes) may be limited. A further constraint is that the encoding format must fully support the interoperability requirements. Thus, encoding of the biometric characteristic with a two-dimensional digital representation of, e.g., a fingerprint image, face image, or iris image is a prominent format structure for many applications. The image itself is stored in standardized formats that allow high compression ratio. Facial images are stored according to JPEG, JPEG2000. For fingerprint images a Wavelet Scalar Quantization (WSQ) has been proven to be a highly efficient encoding. It can be shown that a 300 kbyte image can be compressed to a 10 KByte WSQ file without compromising the biometric performance [4]. Compression formats such as JPEG2000 furthermore can encode a specific region of interest in higher quality using limited compression, and more aggressively compress the remainder background image. A good example is the encoding of the iris in high resolution, while all other areas of the image such as the lids may essentially be masked out. In such a case, images can be compressed down to 2.5 KByte and still yield an acceptable performance [5].

Nonetheless SmartCard-based systems such as the European Citizen Card [6] or the U.S. PIV Card [7] not only require further reduction of the format size but also a good computational preparation of the comparison step especially in environments with low computational power. Comparison-On-Card is an efficient concept to realize privacy protection: The relevant concept is that the biometric reference is not disclosed to the potentially untrusted recognition system. Hence the fresh recognition sample is provided to the card and comparison and decision are performed by the trusted SmartCards. Samples are encoded in a template

format, as a vector of individual features that were extracted from the captured biometric sample. This process is quite transparent as for example in the fingerprint analysis: The essential features of a fingerprint are minutia locations (ridge endings and ridge bifurcations) and directions and potentially extended data such as ridge count information between minutia points. This data is relevant information for almost every fingerprint comparison subsystem and standardizing a minutia format was a straightforward process [8].

These feature-based format standards encode only the structured information – none of the various concepts and algorithms that extract minutia points has been included in the standardization work. Many approaches for these tasks have been published in the academic literature; nevertheless, solutions in products are considered as intellectual property of the suppliers and therefore usually not disclosed.

Furthermore, it became necessary to cope with different cultures in identifying minutia points. Thus minutia definitions based on ridge ending versus definitions based on valley skeleton bifurcations became sub-types of the standard. While these ambiguities cover the variety of approaches of industrial implementations, an impressive interoperability can still be achieved, as it was proven in two independent studies [9, 10].

Requirements from biometric recognition applications are quite diverse: Some applications are tuned on high biometric performance (low error rates) in an identification scenario. Other applications are tuned to operate with a low capacity token in a verification scenario. Where database systems are designed, the record format sub-type is the appropriate encoding. In other applications the token capacity may be extremely limited and thus the card format sub-type that exists in ISO/IEC IS 19794 for the fingerprint data formats in Part 2, 3 and 8 is the adequate encoding. Other parts such as 19794-10, which specifies the encoding of the hand silhouette, have been designed to serve implementations that are constrained by storage space. In general the concept of compact encoding with the card format is to reduce the data size of a BDB down to its limits. This can be achieved when necessary parameters in the metadata are fixed to standard values, which makes it obsolete to store the header information along with each individual record.

For all data interchange formats it is essential to store along with the representation of the biometric characteristic essential information (metadata) on the capturing processing and the generation of the sample. Note that in the case of the card format sub-type fixed values may be required as discussed above. Metadata that is stored along with the biometric data (the biometric sample at any stage of processing) includes information such as size and resolution of the image and (e.g., fingerprint image, face image) but also relevant data that impacted the data capturing process: Examples for such metadata are the Capture Device Type ID, that identifies uniquely the device that was used for the acquisition of the biometric sample and also the impression type of a fingerprint sample, which could be a plain live scan, a rolled live scan, non-live scan or stemming from a swipe sensor. Furthermore, the quality of the biometric sample is an essential information that must be encoded in the metadata. In general, an overall assessment of the sample quality is stored on a scale from 0 to 100, while some formats allow additional local quality assessment such as the fingerprint zonal quality data or minutia quality in various fingerprint encoding standards [8, 11]. The rationale behind this quality recording is to provide information that might weigh into a recapture decision, or to drive a failure to acquire decision. A biometric system may need to exercise quality control on biometric samples, especially enrollment, to assure strong performance, especially for identification systems. Furthermore, multimodal comparison solutions should utilize quality to weigh the decisions from the various comparison subsystem to improve biometric performance. Details on how to combine and fuse different information channels can be found in the ISO technical report on multibiometric fusion [12]. A local quality assessment may also be very meaningful as environmental factors (such as different pressure, moisture, or sweat may locally degrade the image quality of a fingerprint) and thus degrade biometric performance.

In general the metadata in an ISO data interchange format is subdivided into information related to the entire record which is stored in the general header and specific information related to one individual view, which is stored in the view header. The existence of multiple views is of course dependent on the application and the respective modality used. In the case of a fingerprint recognition system it is a common approach, in order to achieve a higher recognition performance, to store multiple views such as right and left index finger together as separate views in one BDB.

The general structure of ISO data interchange format standards is:

1. General header
2. View 1 (mandatory)
   a. View header
   b. View data
3. Views 2-N (optional)
   a. View header
   b. View data

This structure is not yet implemented in all Parts of ISO/IEC 19794, but harmonization in this regard is expected in the revision process of these standards.

Common elements of the general header are the format identifier, the version number of the standard, the length of the record, Capture Device ID, the number of views in the record, and other complementary information.

Elements of the view header are dependent on the modality in use. Typical represented information for a biometric fingerprint sample includes the finger position (right thumb, right index finger, . . ., left index finger, . . ., left little finger), the view number (in the record), the impression type (live-scan plain, live-scan rolled, etc.), finger quality, and number of minutia.

Often, the mere specification for the encoding of the biometric data and the metadata is not enough to assure interoperability. For some biometric modalities, the context of the capture process is also important, and best practices for the capture procedures of the biometric characteristics are described in the standards. The capture of face images suitable for biometric comparison is described in an amendment to ISO/IEC IS 19794-5 [13]. This amendment provides suitable constraints for illumination, backgrounds, and how to avoid shadows on the subject's face. Other standards, such as the iris standard ISO/IEC IS 19794-6 include such information in an annex of the base standard [14].

Published Standards

After the international subcommittee for biometric standardization, SC37, was founded in 2002 [15]. The first standards were already published after an extremely short preparation period in summer 2005.

Standardization in the field of information technology is pursuit by a Joint Technical Committee (JTC) formed by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). An important part of the JTC1 SC37 subcommittee's activities is the definition of data interchange formats in its Working Group 3 (WG3) as described in the previous section. WG3 has, over its first years of work concentrated on the development of the ISO 19794 family, which includes currently the following 13 parts:

Part 1: Framework (IS)
Part 2: Finger minutiae data (IS)
Part 3: Finger pattern spectral data (IS)
Part 4: Finger image data (IS)
Part 5: Face image data (IS)
Part 6: Iris image data (IS)
Part 7: Signature/Sign time series data (IS)
Part 8: Finger pattern skeletal data (IS)
Part 9: Vascular image data (IS)
Part 10: Hand geometry silhouette data (IS)
Part 11: Signature/Sign processed dynamic data (WD)
Part 12: - void -
Part 13: Voice data (WD)
Part 14: DNA data (WD)

The first part includes relevant information that is common to all subsequent modality specific parts such as an introduction of the layered set of SC37 standards and an illustration of a general biometric system with a description of its functional sub-systems namely the capture device, signal processing sub-system, data storage sub-system, comparison sub-system, and decision sub-system. Furthermore, this framework part illustrates the functions of a biometric system such as enrolment, verification, and identification, and explains the widely use context of biometric data interchange formats in the CBEFF structure.

Part 2–Part 14 then detail the specification and provide modality related data interchange formats for both image interchange and template interchange on feature level. The 19794-family gained relevance, as the International Civic Aviation Organization (ICAO) adopted image-based representations for finger, face, and iris for storage of biometric references in Electronic Passports [13, 14, 16]. Thus the corresponding ICAO standard 9303 includes a normative reference to ISO 19794 [17].

Another relevant standard for the global exchange of biometric data has been developed by the National Institute of Standards and Technology (NIST) as American National Standard [18]. This data format

is the de-facto standard for the interchange of fingerprint and facial information for forensic purposes among criminal police institutions. It is also intended to be used in identification or verification processes. This standard supports fingerprint images, fingerprint minutia, iris images, face images, as well as support for any CBEFF encapsulated biometric data.

The American and Japanese standardization committees are developing national standards in parallel to the SC37 international standards. Many of the projects inside SC37 had been initiated by and received significant support from national standard developments. However with the full constitution of SC37 as one of the most active and productive committees inside the JTC1 many national standardization committees – and essentially all European countries – have stopped the development of pure national standards. Most of the available resources are now focused on and invested in the development and procurement of international standards with the JTC1.

Interoperability and Future Needs

With the current set of data format standards open biometric systems can be developed, which can provide interoperability among suppliers. However, as the prime purpose of a biometric system is to achieve a good recognition performance, a core objective is to achieve a good interoperability performance, e.g., the performance associated with the use of a generator and comparison subsystems from different suppliers. This goal of good interoperability performance can be achieved when conformance of each supplier to the data form standard is reached. The concept of conformance testing supports customers and suppliers. A conformance testing protocol verifies that data records generated by an implementation are compliant to the standard. Testing can be subdivided in three levels:

1. Data format conformance: proof that data fields specified in a data format standard do exist and are filled in a consistent manner. The result of this test indicates whether all the fields are included and values in those fields are in the defined range. This check is conducted on a field-by-field and byte-by-byte operation and is often referred to as "Level 1 conformance testing."
2. Internal consistency checking: In the second level of conformance testing the data record is tested for internal consistency, such as relating values from one field of the record to other parts or fields of the record are conformant. This test is often referred to as "Level 2 conformance testing."
3. Semantic conformance: In the third level of conformance testing the values in the data fields are investigated whether or not they are faithful representation of the biometric characteristic, e.g., for a fingerprint image whether minutia points identified are indeed bifurcation or end points of papillary ridges. The test requires standardized sample data on the one hand and elaborated semantic conformance tests, that are yet not developed.

Along with the definition of conformance testing standards, the standardization of sample quality standards is the most important and pressing work to be solved in SC37. The standardization of quality scores is important as it allows for increased interoperability of reference data. The system that utilizes a biometric reference enrolled under a different quality policy may still be able to leverage that reference if it can understand and make use of the quality information relevant to that biometric reference. Thus, the quality standards and technical reports provide guidance to assure interoperability. The technical reports provide guidance about what is relevant to comparability that should be measured for a given biometric characteristic. Currently, quality standardization exists for an overall framework, along with guidance for fingerprint images and face images.

The SC37 standards community has also initiated the revision projects for the 19794 standards. This revision process will not only enable further harmonization of all 19794-parts under one framework, but also respect technology innovations and discuss options, whether or not for future-proof usage of the standard the encoding of the data fields should support an XML encoding.

Related Entries

► Biometric System Design, Overview
► Common Biometric Exchange Formats Framework Standardization
► Conformance Testing for Biometric Data Interchange Formats, Standardization of
► Face Image Data Interchange Format
► Finger Data Interchange Format, International Standardization

▶ Hand Data Interchange Format, Standardization
▶ International Standardization of Biometrics
▶ Iris Image Data Interchange Formats, Standardization
▶ Speaker Recognition, Overview
▶ Vascular Image Data Format, Standardization

References

1. International Civil Aviation Organization TAG 15 MRTD/NTWG: Biometrics Deployment of Machine Readable Travel Documents. Version 2.0. ICAO (2004)
2. EU-Council Regulation No 2252/2004 – of 13 December 2004 on standards for security features and biometrics in passports and travel documents issued by Member States
3. International Standards ISO/IEC 19785-1: Information technology-Common Biometric Exchange Formats Framework - Part 1: Data element specification (2006)
4. Funk, F., Arnold, M., Busch, C., Munde, A.: Evaluation of image compression algorithms for Fingerprint and face recognition systems. In Proceedings from the Sixth IEEE Systems, Man Cybernetics (SMC): Information Assurance Workshop of Systems, Man and Cybernetics, IEEE Computer Society, pp. 72–78. West Point, NY, USA (2005)
5. Daugman, J., Downing, C.: Effect of severe image compression on iris recognition performance. Technical Report, no 685, Universion of Cambridge, ISSN 1476-2986, (2007)
6. European Citizen Card: CEN TC 224 WG 15: Identification card systems
7. National Institute of Standards and Technology, Biometric Data Specification for Personal Identity Verification. NIST Special Publication 800-76-1 http://csrc.nist.gov/publications/nistpubs/800-76-1/SP800-76-1_012407.pdf (accessed in 2007)
8. International Standards ISO/IEC IS 19794-2: Information technology – Biometric data interchange formats – Part 2: Finger minutia data (2005)
9. National Institute of Standards and Technology: MINEX – Performance and Interoperability of the INCITS 378 Fingerprint Template. http://fingerprint.nist.gov/minex04/minex_report.pdf (accessed in 2006)
10. The Minutia Template Interoperability Testing Project MTIT, http://www.mtitproject.com (accessed in 2007)
11. International Standards ISO/IEC IS 19794-8: Information technology – Biometric data interchange formats – Part 8: Finger pattern skeletal data (2006)
12. International Standards ISO/IEC TR 24722, Multimodal and Other Multibiometric Fusion (2007)
13. International Standards ISO/IEC IS 19794-5: Information technology – Biometric data interchange formats – Part 5: Face image data (2005)
14. International Standards ISO/IEC IS 19794-6: Information technology - Biometric data interchange formats - Part 6: Iris image data (2005)
15. ISO/IEC JTC 1 on Information Technology Subcommittee 37 on Biometrics (http://www.jtc1.org)
16. International Standards ISO/IEC IS 19794-4: Information technology – Biometric data interchange formats – Part 4: Finger image data (2005)
17. International Civil Aviation Organization: Supplement to Doc9303-part 1 Sixth edn. (2006)
18. National Institute of Standards and Technology: American National Standard for Information Systems - Data Format for the Interchange of Fingerprint, Facial, & Other Biometric Information – Part 1. Published as NIST Special Publication pp. 500–271 (May 2007)

Biometric Data Interchange Record (BDIR)

A BDIR is a data package containing biometric data that claims to be in the form prescribed by a specific biometric data interchange format standard.

▶ Conformance Testing for Biometric Data Interchange Formats, Standardization of

Biometric Data Interchange Standard

Biometric Data Interchange Standard is a published documentary specification of a data record for clear exchange of subject's biometric data between two parties.

▶ Biometric Sample Quality, Standardization
▶ Face Image Data Interchange Formats, Standardization
▶ Interoperable Performance

Biometric Decision Time, and the External Operation Time

▶ Operational Times

Biometric Devices

▶ Biometric Sensor and Device, Overview

Biometric Encryption

▶ Encryption, Biometric

Biometric Engines

▶ Biometric Algorithms

Biometric Features

Biometric features are the information extracted from biometric samples which can be used for comparison with a biometric reference. For example, characteristic measures extracted from a face photograph such as eye distance or nose size etc. The aim of the extraction of biometric features from a biometric sample is to remove superfluous information which does not contribute to biometric recognition. This enables a fast comparison and an improved biometric performance, and may have privacy advantages.

▶ Vascular Biometrics Image Format, Standardization

Biometric Fraud Reduction

▶ Fraud Reduction, Overview
▶ Fraud Reduction, Applications

Biometric Front End

▶ Biometric Sample Acquisition

Biometric Fusion

▶ Multibiometrics

Biometric Fusion Standardization

▶ Multibiometrics and Data Fusion, Standardization

Biometric Fusion, Rank-Level

▶ Fusion, Rank-Level

Biometric Header, Standards

▶ Common Biometric Exchange Formats Framework Standardization

Biometric Identity Assurance Services

Matthew Swayze
Principal Technical Consultant,
Daon, Inc.,
Reston, VA, USA

Synonym

BIAS

Definition

Biometric Identity Assurance Services, or BIAS, is a collaborative standards project between the International Committee for Information Technology Standards (INCITS), Technical Committee M1 – Biometrics and the Organization for the Advancement of Structured Information Standards (OASIS). BIAS provides an open framework for deploying and invoking biometric-based identity assurance capabilities that can be readily accessed using services-based frameworks. BIAS services provide basic biometric identity assurance functionality as modular and independent operations that can be assembled in many different ways to perform and support a variety of business processes.

Introduction

In reviewing the current biometric-related standards portfolio and ► service-oriented architecture (SOA) references, it became apparent that a gap exists in the availability of standards related to biometric services. There are several existing biometric-related standards describing how to format either biometric data specifically or transactions containing identity information (including biometric information) for use in a particular application domain. However, these standards do not readily fit into an SOA. As enterprise architectures are increasingly built on SOA models and standards, biometric applications, such as those that perform biometric capture functions, require a consistent set of services to access other biometric-based resources. In this context, a biometric resource could be a database with biometric information, a one-to-many search engine, or a system that performs one-to-one verifications. BIAS seeks to fill the gap by standardizing a set of biometrics-based identity assurance capabilities that applications can invoke remotely across a services-oriented framework in order to access these biometric resources.

Scope

Although focused on biometrics, BIAS recognizes that there are nonbiometric elements to an identity. While the services have been built around biometric-related operations, nonbiometric information can be referenced in several of the service calls. BIAS services do not prescribe or preclude the use of any specific biometric type. BIAS is primarily focused on remote service invocations, and therefore, it does not deal directly with any local biometric devices. Recognizing the need for vendor independence, BIAS attempts to be technology, framework, and application domain independent.

BIAS establishes an industry-standard set of predefined and reusable biometric identity management services that allow applications and systems to be built upon an open-system standard rather than implementing custom one-off solutions for each biometric resource. BIAS defines basic biometric-related business level operations, including associated data definitions, without constraining the application or business logic that implements those operations. The basic BIAS services can be assembled to construct higher level, composite operations that support a variety of business processes.

INCITS and OASIS Collaboration

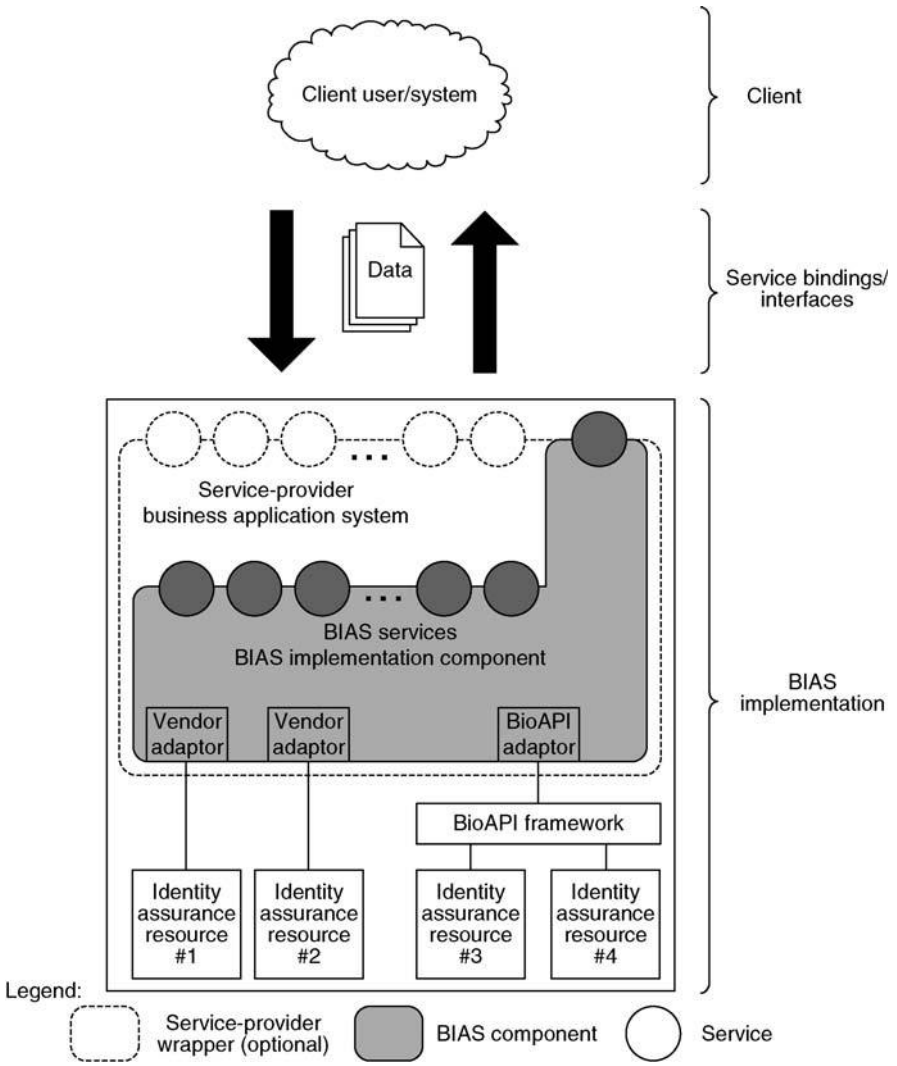
The development of the BIAS standard requires expertise in two distinct technology domains: biometrics, with standards leadership provided by INCITS M1 [1], and service architectures, with standards leadership provided by OASIS [2]. The two groups are collaborating to produce two associated standards. The INCITS M1 standard [3] defines biometric services used for identity assurance, which are invoked over a services-based framework. It is intended to provide a generic set of biometric (and related) functions and associated data definitions to allow remote access to biometric services. The related OASIS standard [4] specifies a set of patterns and bindings for the implementation of BIAS operations (which are defined in the INCITS M1 standard) using Web services and service-oriented methods within XML-based transactional Web services and service-oriented architectures.

Existing standards are available in both fields and many of these standards provide the foundation and underlying capabilities upon which the biometric services depend. The INCITS standard leverages the existing biometric and identity-related standards and formats. The OASIS standard leverages known information exchange and assurance patterns (such as message reliability acknowledgments), and functions (such as repository use and calls) arising in service-oriented systems, and potentially leverages those functions and features that are already embedded in existing SOA methods and standards.

Currently, the INCITS M1 standard has been published as INCITS 442. The OASIS standard, which depends on the INCITS M1 standard, is still in draft form in the OASIS technical committee and is expected to be finalized in 2009.

Architecture

The BIAS architecture consists of the following components: BIAS services (interface definition), BIAS data (schema definition), and BIAS bindings. The BIAS services expose a common set of operations to external requesters of these operations. These requesters may be an external system, a Web application, or an intermediary. The BIAS services themselves are platform and language independent. The BIAS services may be implemented with differing technologies on multiple platforms. For example, OASIS is defining Web services bindings for the BIAS services.



Biometric Identity Assurance Services. **Figure 1** BIAS Application Environment. ITIC. This material is reproduced from INCITS 422-2008 with permission of the American National Standards Institute (ANSI) on behalf of the Information Technology Industry Council (ITIC). No part of this material may be copied or reproduced in any form, electronic retrieval system or otherwise, or made available on the Internet, a public network, by satellite, or otherwise without the prior written consent of the ANSI. Copies of this standard may be purchased from the ANSI, 25 West 43rd Street, New York, NY 10036, (212) 642-4900, http://webstore.ansi.org.

Figure 1 depicts the BIAS services within an application environment. BIAS services provide basic biometric functionality as modular and independent operations that can be publicly exposed directly and/or utilized indirectly in support of a service-provider's own public services.

Services

BIAS defines two categories of services: primitive and aggregate. Primitive services are basic, lower-level operations that are used to request a specific capability. Aggregate services operate at a higher-level, performing a sequence of primitive or other operations in a single request. An example of an aggregate service would be where a one-to-many search (identify), which results in a 'no match,' is immediately followed by the addition of the biometric sample into that search population (enroll).

BIAS provides primitive services for the following areas:

1. Manage subject information: adding or deleting subjects, or associating multiple subjects into a single group.
2. Managing biographic information: adding, updating, deleting, or retrieving biographic information on a particular subject.
3. Managing biometric information: adding, updating, deleting, or retrieving biometric information on a particular subject.
4. Biometric searching/processing: performing biometric one-to-one or one-to-many searches, checking biometric quality, performing biometric fusion, or transforming biometric data from one format to another.

BIAS also defines several aggregate services. The intent of BIAS is to standardize the service request; organizational business rules will determine how the service is actually implemented. The standard aggregate services include Enroll, Identify, Verify, and Retrieve Information.

Summary

The BIAS standard represents the first collaboration between INCITS M1 and OASIS, bringing these two organizations together to define a set of standardized biometric services that can be invoked within a services-oriented framework. The services are defined at two levels and correspond to basic biometric operations. BIAS is technology and vendor independent, and therefore, it may be implemented with differing technologies on multiple platforms.

References

1. INCITS M1 – Biometrics, http://www.incits.org/tc_home/m1.htm. Last Accessed 02 April, 2009
2. OASIS, http://www.oasis-open.org/home/index.php. Last Accessed 02 April, 2009
3. ANSI INCITS 442-2008, Biometric Identity Assurance Services (BIAS), May 2008, http://www.incits.org. Last Accessed 02 April, 2009
4. OASIS BIAS SOAP Profile (Draft), http://www.oasis-open.org/committees/bias
5. Service-Oriented Architecture: Beyond Web Services, Java Developer's Journal, http://java.sys-con.com/read/44368_p.htm. Accessed Feb, 2006
6. Reference Model for Service-Oriented Architecture 1.0, OASIS, http://www.oasis-open.org/committees/download.php/19679/soa-rm-cs.pdf. Accessed Feb, 2007

Biometric Information Record

▶ Common Biometric Exchange Formats Framework Standardization

Biometric Interchange Formats

▶ Biometric Technical Interface, Standardization

Biometric Interfaces

Catherine J. Tilton
Daon, Reston, VA, USA

Definition

Biometric interfaces comprise the methods by which one biometric system component communicates with

another. These components may be devices, software, or entire systems. Implied in this definition is the exchange of information – generally that of biometric data. Interfaces are key elements of biometric system architecture and design and provide the basis for interoperability.

Introduction

Biometric systems are composed of subsystems and components, the configuration and interrelationship of which describe the system architecture. For the system to function, these components must interact with one another across intra-system interfaces. The system itself may be a part of a larger "system of systems" in which inter-system interfaces also exist. In a biometric (or biometrically enabled) system, the interface involves the exchange of biometric data or the invocation of ▶ biometric services.

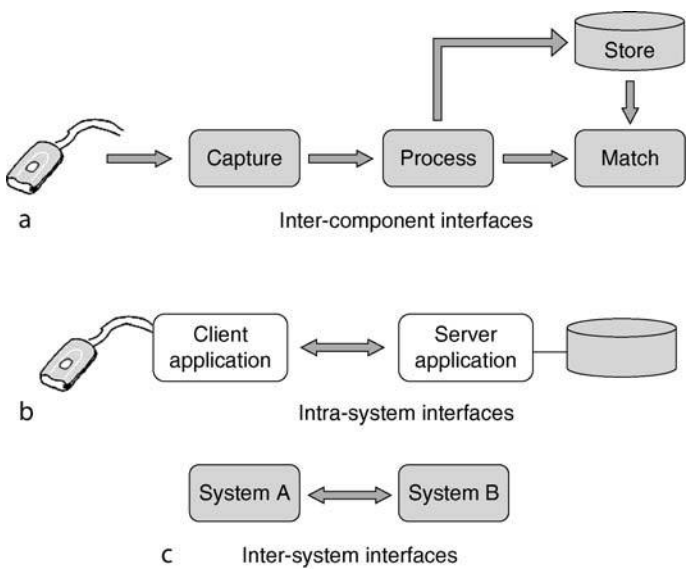
Biometric interfaces exist at a variety of levels – from low-level internal interfaces within a capture device, for instance, to inter-system messaging interfaces, such as between law enforcement systems in different countries (Fig. 1).

The biometric process involves a series of steps including data collection (capture), processing (feature extraction), storage, and matching depending on the operation (i.e., enrollment, verification, or identification). Biometric data may be transferred between components performing these operations or to an application controlling or using the results of the operation. A biometric interface exists whenever biometric data is transferred from one system component to another, internally or externally. The following sections describe data interchange, device interfaces, application programming interface and communications, and messaging interfaces.

Data Interchange

▶ Biometric data may exist in a variety of forms – "raw" biometric sample data captured by a sensor device, partially processed data (e.g., a biometric sample that has undergone a degree of image processing), or a fully processed biometric reference template or recognition sample suitable for matching. Likewise, this data may be formatted and encoded in different ways. An image, for example, can be compressed or uncompressed. Biometric data may exist as a single sample or be packaged together with other like or unlike samples from the same individual. It may exist in a proprietary format or in a standard format, with or without associated metadata.

Whenever data is exchanged between components or systems, the format and encoding of that data must be understood by both the sending and the receiving



Biometric Interfaces. Figure 1 Interface Types and Levels.

entities. This implies that the format information is defined in a document of some type. If both ends of the interface are owned/controlled by the same entity (e.g., a device manufacturer) then the definition may be less formal or be contained within some larger specification. As the relationship between the endpoints becomes more loosely coupled, more formal and rigorous data definitions are needed.

In a closed system, the data format can be whatever works. It can be highly customized and proprietary. In open systems, however, data formats need to be standardized so that they can be understood by a wide variety of producers and consumers of biometric data. Today, data interchange format standards exist for most modalities, although at the raw/image levels. Standard template formats exist only for fingerprint biometrics.

In addition to the biometric data itself, standards exist for encapsulating ("packaging") that data. This includes defined structures, standard metadata headers, and security information. Examples of such standards are the Common Biometric Exchange Formats Framework (CBEFF) and ANSI/NIST ITL1-2007 [1, 2].

More information on data interchange standards can be found in the chapter on Standardization.

Device Interfaces

Biometric sensor devices capture biometric data and sometimes provide additional capabilities to process, store, and/or match it. For an application to integrate a biometric device, an interface to that device must exist and be defined. This includes the physical interface, the communications protocol, and the data/message exchange.

Physical interfaces to biometric devices generally utilize industry standards which define both the physical interface and communications protocols. Because biometric data samples (especially raw data such as images) can be very large, an interface that provides adequate speed and bandwidth is desirable. In the early days of electronic fingerprint scanners, IEEE 1284 parallel interfaces were the norm. Today, the Universal Serial Bus (USB) or IEEE 1394 ("Firewire™") are more commonly used. Some biometric sensors are commodity items such as cameras, microphones, or signature pads.

A common software interface for devices is TWAIN whose purpose is to provide and foster a universal public standard which links applications and image acquisition devices. It supports image acquisition from a scanner, digital camera, or another device and imports it directly into an application. Many commodity devices provide TWAIN-compliant device drivers.

To interface to a biometric device from a software application, operating system (OS) support is required. This is generally accomplished via a "device driver". Most devices provide Windows™ device drivers; however, support for other platforms (such as Linux, Unix, OS2, etc.) is a bit more spotty.

In addition to the device drivers, biometric device manufacturers usually provide software developer kits (SDKs) to control and access the functionality of their device. Applications interface to SDKs via a defined application programming interface (API) as described in the following.

Software Interfaces

Biometric software modules are components that provide a set of biometric functions or capabilities via a software interface. This includes biometric processing and matching algorithms or control of a biometric device. Reusable software packages are called SDKs. Biometric SDKs with standardized interfaces are called biometric service providers (BSPs).

APIs can be either "high level" or "low level". In terms of biometrics, a high level API provides a set of more abstract, generalized functions (e.g., "Enroll") whereas a lower level API provides more specific, atomic functions (e.g., "Capture Fingerprint Image" or "Set Contrast"). The lower the level, the more modality- and even vendor/device-specific it is. An example of a low level biometric API standard is the Speaker Verification API (SVAPI) developed in the mid-90's and championed by Novell [3].

A software application interfaces to an SDK or BSP via an API. The first biometric SDKs appeared in the mid-90's. Most SDK APIs are vendor specific. They are defined by the manufacturer to be highly tailored to the features and capabilities of their product. The advantage of such APIs is that they can be very efficient and provide sophisticated controls. Standard biometric APIs also exist, which define a common interface definition for a category of services. This allows an application to be written once to the standard API and utilize any biometric SDK/BSP that conforms to the standard.

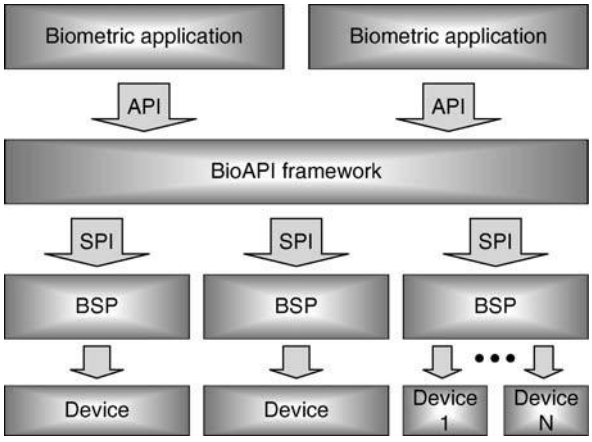
Early APIs were defined using 'C' language constructs. However, more recently the trend is to define object oriented interfaces in terms of Java, .NET, or COM in order to be more easily integrated into object-oriented applications.

The most well known biometric API is the BioAPI. This standard was originally developed by a group of over 100 organizations from industry, government, and academia and published in 2000 as an open systems industry specification [4]. Subsequently, version 1.1 was published as an American National Standard (INCITS 358) and version 2.0 as an international standard (ISO/IEC 19784) [5, 6].

The BioAPI interface defines a set of functions (and associated data structures), including biometric, database, and unit (device) operations, component management functions, utility functions, and data handle, callback, and event operations. High level biometric operations such as enroll, verify, and identify are provided as well as more primitive operations such as capture, create template, process (feature extraction), verify match, identify match, and import. Conformance categories identify which functions and options are required for a given product class.

To perform module management functions, a BioAPI framework component is included as part of the API/SPI (service provider interface) architecture. This allows dynamic insertion and control of BSPs and devices, as well as a discovery mechanism (Fig. 2).

BioAPI is defined as a 'C' interface, though a Java version is in progress. The version 1.1 framework has been ported to Win32, Linux, Solaris, and WinCE platforms and a variety of wrappers (e.g., JNI, C#) have been developed.



Biometric Interfaces. Figure 2 BioAPI Architecture.

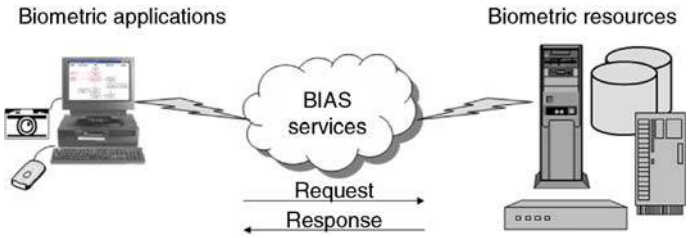
Another "standard" biometric API is BAPI. This API was developed by I/O Software and later licensed to Microsoft who included it in their XP Home Edition as the interface to their fingerprint device. This API originally provided 3-levels of interface – a high level similar to BioAPI, a mid-level, and a lower (device) level interface. BAPI has not been made publicly available or formally standardized.

More recently, the Voice Extensible Markup Language (VoiceXML) was created for creating voice user interfaces that use automatic speech recognition (ASR) and text-to-speech synthesis (TTS). It was developed by the VoiceXML Forum and published by the W3C. "VoiceXML simplifies speech application development by permitting developers to use familiar Web infrastructure, tools and techniques. VoiceXML also enables distributed application design by separating each application's user interaction layer from its service logic." [7] An extension to VoiceXML called Speaker Identification and Verification (SIV) is in progress [8].

In addition to group developed APIs, there have been biometric APIs developed by application and middleware vendors. The latter standardize an interface to their particular product or product line. In this case, the application/middleware vendor defines an interface such that any biometric technology vendor wishing to be integrated (or resold) with that application must conform to the application vendor's API. It may be a biometric specific API or a more general "authentication method" API. While this has been successful to some extent, the drawback is that the technology vendors must provide different flavors of their SDK for each such application, which may become difficult to maintain.

Communications and Messaging Interfaces

When biometric information is passed between systems or subsystems, a communications or messaging interface may be used. This is generally defined in terms of message content and protocol. The best known are those used by the justice community. The FBI's Electronic Fingerprint Transmission Specification (EFTS) and the Interpol Implementation (INT-I) both utilize the ANSI/NIST ITL1-2000 standard to define



Biometric Interfaces. **Figure 3** Biometric Web Services Using BIAS.

transactions (request and response messages) with their respective systems [2, 9, 10] (note that the EFTS and ANSI/NIST standards have recently been revised; however, at the time of this article, they had not yet been implemented. Interpol is expected to follow suit) [11, 12].

ANSI/NIST ITL1-2000/2007 defines the content, format, and units of measurement for electronically encoding and transmitting fingerprint, palmprint, facial/mugshot, and SMT images, and associated biographic information. It consists of a series of "record types", each containing a particular type and format of data. For example, a Type-4 record contains a high resolution grayscale fingerprint image, a Type-9 record contains minutiae data, a Type-10 facial or SMT images, a Type-14 variable-resolution tenprint images, etc. An XML version of the 2007 standard was recently released [13].

EFTS and INT-I define transactions in terms of these records and further define the content of "user-defined fields". For example, EFTS defines a type of transaction (TOT) called a CAR (Criminal Tenprint Submission, Answer Required) that "contains ten rolled and four plain impressions of all ten fingers, as well as information relative to an arrest or to custody or supervisory status and optionally may include up to 4 photos of the subject." [8] It nominally consists of a Type 1 (header), Type 2 (descriptive text), 14 Type-4, and 0-4 Type-10 records.

Services Interfaces

Today's biometric systems are being built upon what is commonly referred to as a "service oriented architecture (SOA)". In an SOA, requesting applications/systems are decoupled from those systems which provide biometric services and allows biometric operations to be invoked and resources to be accessed remotely, usually across an open or closed network, including the internet. These services interfaces may be customized or standardized.

The most often used protocols for such services are XML over Hypertext Transmission Protocol (HTTP) or Simple Object Access Protocol (SOAP) over HTTP. SOAP services are defined in terms of ▶ Web Services Definition Language (WSDL) and frequently utilize a set of existing web services standards. Service providers may post their WSDL to a directory which can be read by potential users or, in closed systems, may be provided directly to known requesters.

A service provider offers a set of remote biometric services such as biometric data storage and retrieval, 1:1 face verification, or 1:N iris or fingerprint search/match. The requester invokes the operation by sending a service request with the associated data to the service provider. The service provider accepts the request, performs the operation, and returns the results as a service response (Fig. 3).

Although today most services interfaces are system specific, a project known as Biometric Identity Assurance Services (BIAS) is in progress to standardize a set of generic biometric Web services. (See BIAS section of the Standardization chapter for more information.)

Summary

Biometric interfaces provide a means to exchange biometric data, perform data transactions, and invoke biometric services. This can occur at several different levels and between different types of biometrics and system components. All biometric interfaces involve transfer of biometric data and must be specified in some way. An interface definition may be proprietary, as is frequently done in closed systems, or standardized. Biometric interfaces are key aspects of the overall biometric system architecture and design.

Related Entries

- ▶ Biometric Sensor and Device, Overview
- ▶ Biometic System Design, Overview
- ▶ Standardization

References

1. Common Biometric Exchange Formats Framework (CBEFF), INCITS 398-2008 and ISO/IEC 19785-1:2006
2. American National Standards Institute and National Bureau of Standards, ANSI/NIST ITL1-2000: Data Format for the Interchange of Fingerprint, Facial, and SMT Information
3. Speaker Verification API (SVAPI), SVAPI Working Group, originally published in 1997 with latest version 2004, http://developer.novell.com/wiki/index.php/SRAPI and SVAPI Source Code. [Note also article by J. Markowitz Introduction to SVAPI at http://www.jmarkowitz.com/downloads.html.]
4. BioAPI Consortium website: http://www.bioapi.org
5. BioAPI Consortium, American National Standards Institute (ANSI) and International Council on Information Technology Standards: The BioAPI Specification, Ver. 1.1, INCITS 358-2002, Feb 2002
6. ISO/IEC 19784-1: Information technology – Biometric application programming Interface – Part 1: BioAPI specification, Ver. 2.0, 1 May 2005
7. VoiceXML Forum website: http://www.voicexml.org
8. W3C website: http://www.w3.org/voice
9. Federal Bureau of Investigation: Electronic Fingerprint Transmission Specification (EFTS), IAFIS-DOC-01078-7.1, May 2, 2005
10. Interpol AFIS Expert Group: Interpol Implementation of ANSI/NIST ITL1-2000, Version No. 4.22b, October 28, 2005
11. American National Standards Institute and National Bureau of Standards, ANSI/NIST ITL-1-2007: Data Format for the interchange of Fingerprint, Facial, and other Biometric Information – Part 1, April 2007
12. Federal Bureau of Investigation: Electronic Biometric Transmission Specification (EBTS), IAFIS-DOC-01078-8.1, November 2008
13. American National Standards Institute and National Bureau of Standards, ANSI/NIST ITL2-2008: Data Format for the Interchange of Fingerprint, Facial, and Other Biometric Information – Part 2: XML version, Aug 2008

Biometric Key Generation

▶ Encryption, Biometric

Biometric Locking

▶ Encryption, Biometric

Biometric Match-on-Card, MOC

▶ On-Card Matching

Biometric Modality

The biometric characteristic which is used in a biometric process is known as biometric modality.

▶ Multibiometrics and Data Fusion, Standardization

Biometric PAC

▶ Access Control, Physical

Biometric Performance Evaluation Standardization

▶ Performance Testing Methodology Standardization

Biometric Quality Evaluation

▶ Biometric Sample Quality

Biometric Quality Standards

► Biometric Sample Quality, Standardization

Biometric Readers

► Access Control, Physical

Biometric Recognition

► Biometrics

Biometric Reference

One or more stored biometric samples, biometric templates or biometric models attributed to a biometric data subject and used for comparison.

► Biometric Data Interchange Format, Standardization

Biometric Registration Authority

► Common Biometric Exchange Formats Framework Standardization

Biometric Sample

See "Biometric data."

► Biometric Interfaces

Biometric Sample Acquisition

Dale Setlak
AuthenTec, Inc., Melbourne, Fl, USA

Synonyms

Biometric data acquisition; Biometric data capture; Biometric front end; Biometric sensing; Fingerprint capture; Fingerprint reading; Fingerprint scan; Image capture; Iris capture; Iris scan

Definition

Biometric sample acquisition is the process of capturing information about a biological attribute of the subject, as it exists within a specific time frame. The objective is to measure data that can be used to derive unique properties of the subject that are stable and repeatable over time and over variations in acquisition conditions.

Typically, the capture process measures a physical property that is affected by the biological characteristic of interest, and converts the measured data into a format that is suitable for analysis – typically a digital electronic format compatible with computerized analysis.

For simplicity, in this discussion, it is assumed that behavioral biometrics are biological attributes that have a temporal dimension and are included in the discussion as such.

In classic biometric systems such as criminology systems, there is a definitive separation (in both time and space) between biometric sample acquisition and the processing and matching of that sample. For example, an arresting officer may collect a suspect's fingerprints at a booking station in the sheriff's office. The fingerprints may then be sent to the FBI for processing and matching against a fingerprint repository. In contrast, real-time biometric ID verification systems, such as those used for login on a laptop computer, do not have that clear separation. In laptop computers, for example, the sample processing and matching will begin operating, while sample acquisition is still in progress. Information from those analyses can then be used to optimize the sample acquisition in real time, significantly improving the overall performance of the system, but blurring the

separation between sample acquisition and the subsequent processes.

Introduction

This article will start out by examining the high level requirements that apply generally to many types of biometric sample acquisition. The biometric sample acquisition process will then be decomposed into its essential elements and each of those discussed briefly. It then examines how each of the essential elements is applied, using the fingerprint ridge pattern as the example biological property, and also examines the real-world implementation embodied in the recently popular fingerprint login systems on laptop computers. The article then reviews some of the new requirements imposed on biometric sample acquisition systems when they become essential elements of the secure, trusted computing, and communication systems that are needed by applications such as mobile commerce and mobile banking.

Generalized Requirements for Biometric Sample Acquisition

The fundamental requirements for the biometric sample acquisition process are driven by the needs of the biometric matching process. At the conceptual level, these requirements boil down to the following two:

- To be able to distinguish a large number of people from each other, a biometric property must contain a large amount of information entropy. In state space terms, the property must have a very large number of distinguishable states. As a result, most biometric characteristics are complex properties represented as arrays of information such as 2- or 3-dimensional images of biological structures (e.g., fingerprints), or segments of time series data (e.g., speech segments). Biometric sample acquisition then becomes the task of making a large number of measurements that have well known interrelationships in space and/or in time, with sufficient resolution and accuracy to develop the required large measurement state space.
- To avoid failing to recognize a previously enrolled person, the biometric matching process needs repeatable detail among all the samples of the biometric property data. The key is minimizing sample variability. Ideally, the biometric sample acquisition system should capture the same biometric property data across the full range of conditions in which it is used. This can become a significant challenge given the wide variability in the biological structures being measured across the human population and the wide range of environmental conditions in which some biometric systems must function.

Sample variability can come from a variety of sources including:

- Intrinsic biological variability
- Environmental variability
- Sample presentation variability
- Biological target contamination
- Acquisition losses, errors, and noise

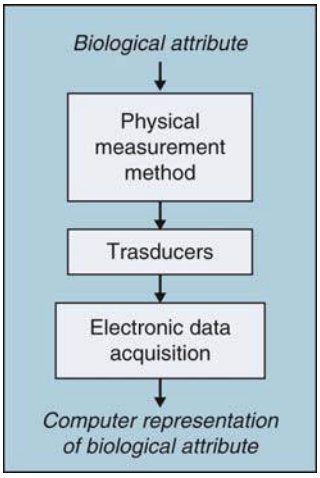
Good biometric sample acquisition systems minimize the effects of these sources of variability.

Process Decomposition

For our discussion here, the sample acquisition process can be decomposed into three parts:

1. The measurement physics
2. The transducers
3. The electronic data acquisition

Figure 1 illustrates this decomposition.



Biometric Sample Acquisition. Figure 1 Biometric sample acquisition process decomposition

The Measurement Physics

Starting with a biological attribute of interest, a physical sensing method is selected, usually involving an energy flow that originates from, or has been modified by the biological attribute to be measured.

Different physical sensing mechanisms may be more or less sensitive to the biological attribute of interest. A key element in selecting the sensing mechanism is the intrinsic signal to noise ratio. A mechanism that has high sensitivity to the attribute of interest and low sensitivity to other influences is likely to have a favorable signal to noise ratio [1].

The Transducers

The energy flow associated with a sensing method may be measurable by several different types of transducers. Transducers convert the energy associated with a physical measurement into a representative electronic signal. Different transducers may be more or less effective in extracting the biological information from the energy flow.

The Electronic Data Acquisition

Electronic data acquisition equipment converts the transducer output signal into a standardized form that can be manipulated by digital computers [2]. This digitized data becomes the input to the feature extraction and pattern matching processes.

The data acquisition process typically involves [3]:

- Generating excitation energy and applying it to the biological structures to be measured and/or to the transducers
- Amplifying the transducer signals
- Multiplexing the signals from a multitude of transducers to a small number of signal processing nodes
- Canceling or filtering noise in the transducer signals
- Time-sampling the transducer signals
- Digitizing the (typically analog) transducer signals
- Assembling the digitized signals into a formatted data stream for delivery to a microprocessor for further processing [4]

An Example Biometric Sample Acquisition Process

For example, select the fingerprint ridge pattern as the biological attribute to be measured.

Example Sensing Physics and Transducers

The fingerprint ridge pattern is able to generate or influence several different types of energy, and hence, may be amenable to several different measurement methods. Each type of energy can be measured by several types of transducers. Designing the biometric sample acquisition system then involves finding the optimum combination of measurement methods and transducer type for the application [5].

Pressure

Fingerprint ridges and valleys can apply different amounts of pressure to a contact surface. A wide variety of transduction methods can detect such spatial pressure variations. These span the range from arrays of tiny nano-switches, to the legacy inkpad and card systems used with fax-machine-like card scanners.

Optical Energy

Fingerprint ridges and valleys differ in their abilities to reflect light, absorb light, and diffuse light. When one of the various forms of optical energy has been applied to the fingerprint region of the skin, camera-like image capture devices can then capture the fingerprint patterns from the resulting light. Figure 2 illustrates the energy conversions involved in a typical optical fingerprint reader.

Electrical Energy

The fingerprint ridge and valley pattern can affect the movement of electrical energy in several different ways, and electrical energy can be measured by several different types of transducers. Arrays of electrical transducers then measure the patterns in the electrical energy

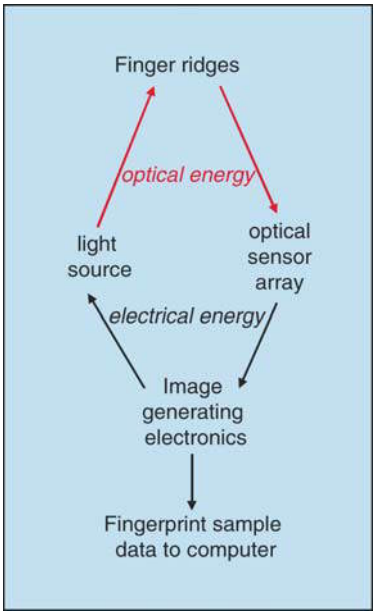
flow to develop a 2-dimensional image of the fingerprint pattern that can be very similar to the images produced by optical measurements.

Acoustical and Thermal Energy

Both acoustical energy and thermal energy propagate more efficiently through the fingerprint ridges than through the air spaces in the valleys between the ridges. Arrays of acoustical and thermal transducers then can detect the pattern of ridges in contact with the array, and generate images similar to those produced by the optical and electrical methods discussed above.

Example Electronic Data Acquisition

Arrays of all the above types of transducers can be fabricated today on the top surfaces of silicon integrated circuits [6]. The transducers can then be connected directly to silicon electronic circuits that perform the data acquisition tasks described in the previous section of this article. The integration of arrays of transducers with data acquisition circuitry on a single silicon chip has reduced the size and cost of biometric sample acquisition systems by a factor of over 100 within the 10 years between 1997 and 2007, enabling a wide variety of new biometric identity verification applications that had previously been cost prohibitive.



Biometric Sample Acquisition. Figure 2 Example of optical fingerprint sample acquisition process

Real-World Implementation –Biometric Sample Acquisition Systems in Widespread Use Today

If you have a laptop computer purchased in 2007 or later, there is a good chance as it has a biometric sample acquisition system built into it – in the form of a small fingerprint sensor integrated into the keyboard. The fingerprint sensor can be used as a convenient alternative to passwords when you logon to your computer, or when you access a password protected website. Figure 3 is a photograph of a laptop computer with a built-in fingerprint sensor.

The fingerprint sensors integrated into laptop computers use tiny bits of electrical energy as discussed above to detect the fingerprint pattern of a finger when you slide your finger across the sensor. There are two types of sensing physics in common use in today's laptops. One type uses electrical energy to measure differences in electrical capacitance between pixels near a fingerprint ridge and pixels near a valley. The other type uses small radio frequency signals to detect the fingerprint shape in the conductive layer of skin just beneath the surface. Both types of sensors are fabricated as silicon devices, with integrated transducers and data acquisition electronics.

Biometric Sample Acquisition. Figure 3 Fingerprint Sensor in Laptop computer

New Requirements for Security in Biometric Sample Acquisition

Biometric sample acquisition systems have begun to take a roll in user identity verification in mobile computing and communication systems. Examples include the previously discussed fingerprint enabled laptops and biometrically secured cellphones as well. This is a different type of role than that played by biometric systems in the forensic and criminology worlds, because these new systems operate in unsupervised and usually insecure situations. This section examines some of the implications of that new role for biometric sample acquisition and the new requirements imposed on biometric sample acquisition by that role.

Using Biometric Data in Trustworthy Identity Verification

While biometric verification is often used as a replacement for passwords, biometric methods when applied to identity verification function more like a handwritten signature and less like a password. This is not surprising, since a handwritten signature is considered as a form of biometric identity verification.

It can be argued that biometric sample data of any kind cannot be considered secret, hence ▶ Trustworthy Biometric Identity Verification in unsupervised situations requires the biometric sample acquisition system to function as a kind of trusted agent [7], essentially certifying (to some reasonable degree of confidence) the validity of the biometric sample that it generates. The role is somewhat analogous to that of a Notary in handwritten signature situations. This new role imposes new requirements on the biometric sample acquisition system that do not exist in the heavily supervised biometric acquisition processes associated with criminology and forensics.

While it is not the intention here to discuss the full scope of trusted biometric identity verification systems, the biometric sample acquisition part of that system inherits certain requirements that can be discussed in this context. Thus, for biometric sample acquisition systems designed to function within unattended identity verification systems, the added requirements include resistance to a number of attack vectors that could be used to falsify the biometric sample that the system delivers.

Trusted Biometric Sample Acquisition Systems

A trusted biometric sample acquisition system inherits atleast the following requirements:

- Resistance to fake biometric target presentation.
- This capability is also called ▶ Biometric Spoof Prevention. It provides an appropriate degree of protection against attacks like the use of a face mask to fool a face recognition system, or movie hero James Bond's use of molded latex rubber finger coverings to fool a fingerprint reader.
- Resistance to acquisition system tampering.
- The requirement here is to prevent an attacker from accessing the internal operation of the biometric sample acquisition system, where he could force it to output different information than it is actually measuring. This requirement may impose hardened packaging requirements on the biometric sample acquisition system.
- Resistance to device/system substitution.
- The system as a whole should be able detect if an alternate device has been substituted for all or any portion of the biometric sample acquisition system. This typically imposes cryptographic capabilities on the biometric sample acquisition system.
- Resistance to communications attacks (e.g., man-in-the-middle, and replay).
- The acquired biometric sample must be securely delivered to the subsequent processing stages either by a physically inaccessible data channel or by cryptographic methods.

All these requirements are designed to enhance the trustworthiness of the biometric sample capture event. When a trusted biometric sample acquisition system is integrated into an overall trusted biometric system (e.g., a ▶ sealed local biometric identity verification system), unsupervised biometric identity verification can be performed with reasonable levels of confidence, without concern that biometric properties are intrinsically not secret.

Related Entries

▶ Biometric Applications, Overview
▶ Biometrics, Overview

► Biometrtic Sensor and Device, Overview
► Security and Liveness, Overview

References

1. Fraden, J.: Handbook of Modern Sensors – Physics, Designs, and Applications, 3rd edn. Springer, Heidelberg (2004)
2. Austerlitz, H.: Data Acquisition techniques using PCs. Academic, London (2003)
3. Webster, J.G.: The Measurement Instrumentation And Sensors Handbook. CRC, Boca Raton (1998)
4. Ball, S.: Analog Interfacing to Embedded Microprocessor Systems, 2nd edn. Newnes, Oxford (2003)
5. Ratha, N., Bolle, R.: Automatic Fingerprint Recognition Systems. Springer, Heidelberg (2003)
6. Wilson, J.S.: Sensor Technology Handbook. Newnes, Elsevier, Oxford, UK (2005)
7. Pearson, S.: Trusted Computing Platforms: TCPA Technology in Context; HP Professional Series. Prentice Hall, New Jersey (2003)

Biometric Sample Quality

Elham Tabassi, Patrick Grother
National Institute of Standards and Technology, MD, USA

Synonyms

Biometric quality evaluation; Performance of biometric quality measures

Definition

The intrinsic characteristic of a biometric signal may be used to determine its suitability for further processing by the biometric system or assess its conformance to preestablished standards. The quality score of a biometric sample signal is a scalar summary of the sample's quality.

Quality measurement algorithm is regarded as a black box that converts an input sample to an output scalar. Evaluation is done by quantifying the association between those values and observed matching results. For verification, these would be the false match and nonmatch rates. For identification, the matching results would usually be false match and nonmatch rates [1], but these may be augmented with rank and candidate-list length criteria. For a quality algorithm to be effective, an increase in false match and false nonmatch rates is expected as quality degrades.

Introduction

Biometric quality measurement algorithms are increasingly deployed in operational biometric systems [2, 3], and there is now international consensus in industry [4], academia [5], and government [6] that a statement of a biometric sample's quality should be related to its recognition performance. That is, a quality measurement algorithm takes a signal or image, $\mathbf{x}$, and produces a scalar, $q = Q(\mathbf{x})$, which is predictive of error rates associated with the verification or identification of that sample. This chapter formalizes this concept and advances methods to quantify whether a quality measurement algorithm (QMA) is actually effective.

What is meant by quality? Broadly a sample should be of good quality if it is suitable for automated matching. This viewpoint may be distinct from the human conception of quality. If, for example, an observer sees a fingerprint with clear ridges, low noise, and good contrast then he might reasonably say it is of good quality. However, if the image contains few minutiae, then a minutiae-based matcher would underperform. Likewise, if a human judges a face image to be sharp, but a face recognition algorithm benefits from slight blurring of the image then the human statement of quality is inappropriate. Thus, the term quality is not used here to refer to the ► fidelity of the sample, but instead to the ► utility of the sample to an automated system. The assertion that performance is ultimately the most relevant goal of a biometric system implies that a quality algorithm should be designed to reflect the sensitivities of the matching algorithm. For fingerprint minutiae algorithms, this could be the ease with which minutiae are detected. For face algorithms, it might include how readily the eyes are located.

Quality evaluation methods should not rely on the manual annotation of a data set because this is impractical for all but small datasets, not least because human examiners will disagree in this respect. The virtue of relating quality to performance is that matching trials can be automated and conducted in bulk. The essay notes further that quality algorithms that relate to

human perception of a sample, quantify performance only as much as the sensitivities of the human visual system are the same as those of a biometric matcher.

One further point is that performance related quality evaluation is agnostic on the underlying technology: it would be improper to force a fingerprint quality algorithm to produce low quality values for an image with few minutia when the target matching algorithm is nonminutia based, as is the case for pattern based methods [7].

Evaluation of quality measurement algorithms should be preferably done in large scale offline trials, which offer repeatable, statistically robust means of evaluating core algorithmic capability.

Prior work on quality evaluation, and of sample quality analysis generally, is limited. Quality measurement naturally lags recognition algorithm development, but has emerged as it realized that biometric systems fail on certain pathological samples. Alonso et al. [8] reviewed five algorithms and compared the distributions of the algorithms' quality assignments, with the result that most of the algorithms behave similarly. Finer grained aspects of sample quality can be addressed. For instance, Lim et al. [9] trained a fingerprint quality system to predict the accuracy of minutia detection. However, such methods rely on the manual annotation of a data set, which as stated above is impractical.

Properties of a Quality Measure

This section gives needed background material, including terms, definitions, and data elements, to support quantifying the performance of a quality algorithm. Throughout this chapter, low quality values are used to indicate poor sample properties.

Consider a data set D containing two samples, $d_i^{(1)}$ and $d_i^{(2)}$ collected from each of $i = 1, \ldots, N$ individuals. The first sample can be regarded as an enrollment image, the second as a user sample collected later for verification or identification purposes. Suppose that a quality algorithm Q can be run on the ith enrollment sample to produce a quality value

$$q_i^{(1)} = Q(d_i^{(1)}), \tag{1}$$

and likewise for the authentication (use-phase) sample

$$q_i^{(2)} = Q(d_i^{(2)}). \tag{2}$$

Thus, it has been suggested that these qualities are scalars, as opposed to vectors for example. Operationally, the requirement for a scalar is not necessary: a vector could be stored and used by some application. The fact that quality has historically been conceived of as scalar is a widely manifested restriction. For example, BioAPI [10] has a signed single byte value, BioAPI_QUALITY; and the headers of the ISO/IEC biometric data interchange format standards [11] have five-byte fields for quality with only one byte allocated for quality score. This chapter does not further address the issue of vector quality quantities other than to say that they could be used to specifically direct re-acquisition attempts (e.g., camera settings), and if considered, their practical use would require application of a discriminant function.

The discussion now formalizes the premise that biometric quality measures should predict performance. A formal statement of such requires an appropriate, relevant, and tractable definition of performance. Consider K verification algorithms, V_k, that compare pairs of samples (or templates derived from them) to produce match (i.e., genuine) similarity scores

$$s_{ii}^{(k)} = V_k(d_i^{(1)}, d_i^{(2)}), \tag{3}$$

and similarly nonmatch (impostor) scores

$$s_{ij}^{(k)} = V_k(d_i^{(1)}, d_j^{(2)}) \quad i \neq j. \tag{4}$$

Now, to posit that two quality values can be used to produce an estimate of the genuine similarity score that matcher k would produce on two samples

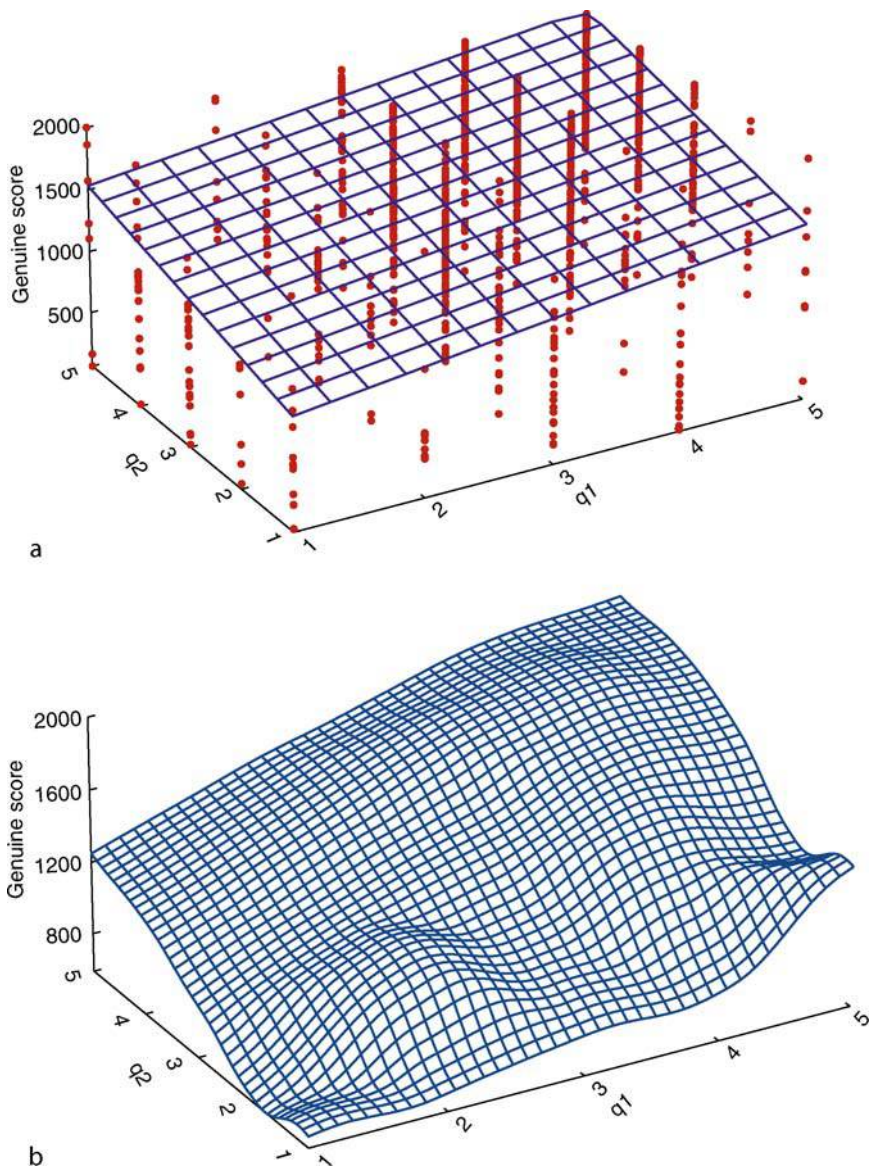
$$s_{ii}^{(k)} = P(q_i^{(1)}, q_i^{(2)}) + \epsilon_{ii}^{(k)}, \tag{5}$$

where the function P is some predictor of a matcher k's similarity scores, and ϵ_{ii} is the error in doing so for the ith score. Substituting (1) gives

$$s_{ii}^{(k)} = P(Q(d_i^{(1)}), Q(d_i^{(2)})) + \epsilon_{ii}^{(k)}, \tag{6}$$

and it becomes clear that together P and Q would be perfect imitators of the matcher V_k in (3), if it was not necessary to apply Q to the samples separately. This separation is usually a necessary condition for a quality algorithm to be useful because at least half of the time (i.e., enrollment) only one sample is available. The obvious consequence of this formulation is that it is inevitable that quality values will imprecisely map to similarity scores, i.e., there will be scatter of the known

scores, s_{ii}, for the known qualities $q_i^{(1)}$ and $q_i^{(2)}$. For example, Fig. 1 shows the raw similarity scores from a commercial fingerprint matcher versus the transformed integer quality scores from NIST fingerprint image quality (NFIQ) algorithm [6, 12], where NFIQ native scores are mapped to $Q = 6 - \text{NFIQ}$ (so that higher quality values indicate good "quality"). Figure 1(a) also includes a least squares linear fit, and Fig. 1(b) shows a cubic spline fit of the same data. Both trend in the correct direction: worse quality gives lower similarity scores. Even though the residuals in the spline fit are smaller than those for the linear, they are still not small. Indeed even with a function of arbitrarily high order, it will not be possible to fit the observed scores perfectly if quality values are discrete (as they are for NFIQ). By including the two fits of the raw data, it is not asserted that scores should be linearly related to the two quality values (and certainly not locally cubic). Accordingly, it is concluded that it is unrealistic to require quality measures to be linear predictors of the similarity scores; instead, the scores should be a monotonic function (higher quality samples give higher scores).



Biometric Sample Quality. Figure 1 Dependence of raw genuine scores on the transformed NFIQ qualities of the two input samples.

Evaluation

Quality measurement algorithms are designed to target application-specific performance variables. For verification, these would be the false match rate (FMR) and false nonmatch rate (FNMR). For identification, the metrics would usually be FNMR and FMR [1], but these may be augmented with rank and candidate-list length criteria. Closed-set identification is operationally rare, and is not considered here.

Verification is a positive application, which means samples are captured overtly from users who are motivated to submit high quality samples. For this scenario, the relevant performance metric is the false nonmatch rate (FNMR) for genuine users because two high quality samples from the same individual should produce a high score. For FMR, it should be remembered that false matches should occur only when samples are biometrically similar (with regard to a matcher) as for example when identical twins' faces are matched. So, high quality images should give very low impostor scores, but low quality images should also produce low scores. Indeed, it is an undesirable trait for a matching algorithm to produce high impostor scores from low quality samples. In such situations, quality measurement should be used to preempt submission of a deliberately poor sample.

For identification, FNMR is of primary interest. It is the fraction of enrollee searches that do not yield the matching entry on the candidate list. At a fixed threshold, FNMR is usually considered independent of the size of the enrolled population because it is simply dependent on one-to-one genuine scores. However, because impostor acceptance, as quantified by FMR, is a major problem in identification systems, it is necessary to ascertain whether low or high quality samples tend to cause false matches.

For a quality algorithm to be effective, an increase in FNMR and FMR is expected as quality degrades. The plots in Fig. 2 shows the relationship of transformed NFIQ quality levels to FNMR and FMR. Figure 2(a) and 2(c) are boxplots of the raw genuine and impostor scores for each of the five NFIQ quality levels. The scores were obtained by applying a commercial fingerprint matcher to left and right index finger impressions of 34,800 subjects. Also shown are boxplots of FNMR and FMR. The result, that the two error rates decrease as quality improves, is expected and beneficial. The FMR shows a much smaller decline. The nonoverlap of the notches in plots of Fig. 2(a) and 2(b) demonstrates "strong evidence" that the medians of the quality levels differ [13]. If the QMA had more finely quantized its output, to $L > 5$ levels, this separation would eventually disappear. This issue is discussed further in section "Measuring Separation of Genuine and Impostor Distributions".

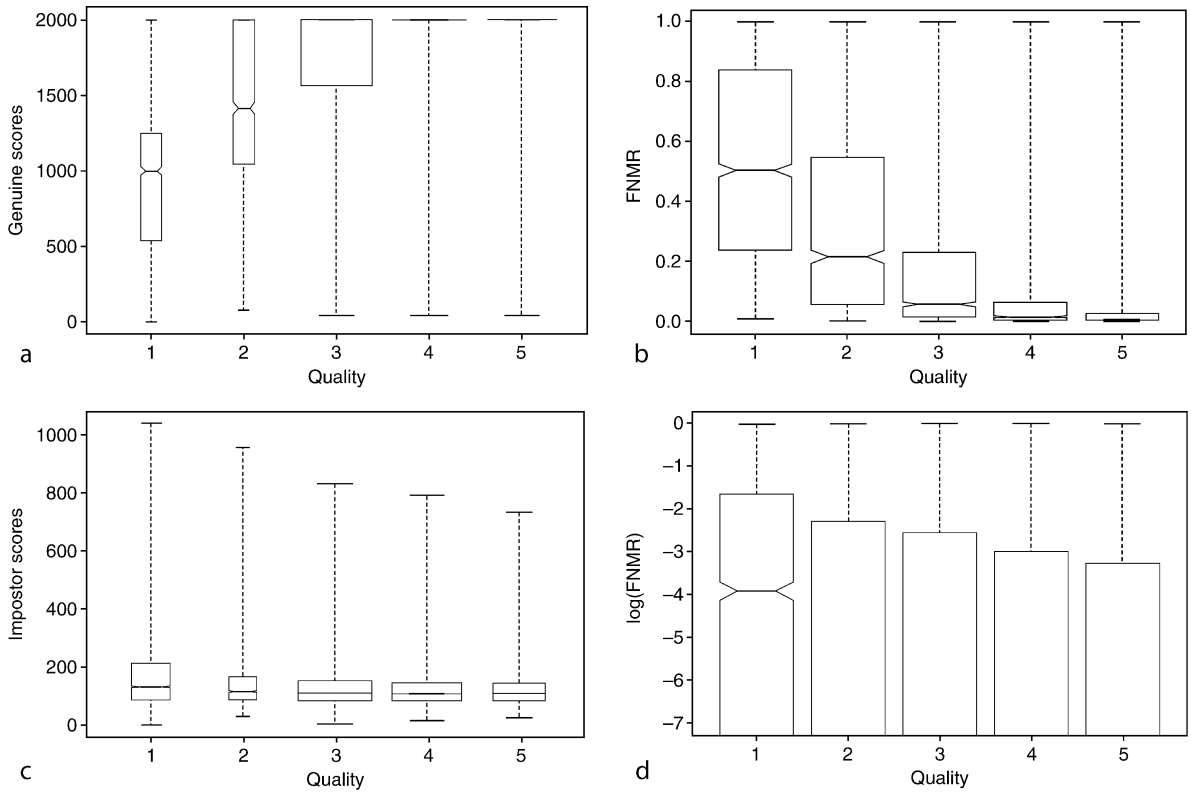
Rank-Ordered Detection Error Tradeoff Characteristics

A quality algorithm is useful, if it can at least give an ordered indication of an eventual performance. For example, for L discrete quality levels there should notionally be L DET characteristics. In the studies that have evaluated performance measures [1, 5, 12, 14, 15, 16], DET's are the primary metric. It is recognized that DET's are widely understood, even expected, but note three problems with their use: being parametric in threshold, t, they do not show the dependence of FNMR (or FMR) with quality at fixed t, they are used without a test of the significance of the separation of L levels; and partitioning of the data for their computation is under-reported and nonstandardized.

This chapter examines three methods for the quality-ranked DET computation. All three use N paired matching images with integer qualities $q_i^{(1)}$ and $q_i^{(2)}$ on the range $[1, L]$. Associated with these are N genuine similarity scores, s_{ii}, and up to $N(N-1)$ impostor scores, s_{ij} where $i \neq j$, obtained from some matching algorithm. All three methods compute a DET characteristic for each quality level k. For all thresholds s, the DET is a plot of $\mathrm{FNMR}(s) = M(s)$ versus $\mathrm{FMR}(s) = 1 - N(s)$, where the empirical cumulative distribution functions $M(s)$ and $N(s)$ are computed, respectively, from sets of genuine and impostor scores. The three methods of partitioning differ in the contents of these two sets. The simplest case uses scores obtained by comparing authentication and enrollment samples whose qualities are both k. This procedure (see for example, [17]) is common but overly simplistic. By plotting

$$\mathrm{FNMR}(s,k) = \frac{\left|\left\{s_{ii}:\ s_{ii} \le s,\ \ q_i^{(1)} = q_i^{(2)} = k\right\}\right|}{\left|\left\{s_{ii}:\ s_{ii} < \infty,\ q_i^{(1)} = q_i^{(2)} = k\right\}\right|},$$

$$\mathrm{FMR}(s,k) = \frac{\left|\left\{s_{ij}:\ s_{ij} > s,\ \ q_i^{(1)} = q_j^{(2)} = k,\ i \neq j\right\}\right|}{\left|\left\{s_{ij}:\ s_{ij} > -\infty,\ q_i^{(1)} = q_j^{(2)} = k,\ i \neq j\right\}\right|}, \tag{7}$$



Biometric Sample Quality. **Figure 2** Boxplots of genuine scores, FNMR, impostor scores, and FMR for each of five transformed NFIQ quality levels for scores from a commercial matcher. Each quality bin, q, contains scores from comparisons of enrollment images with quality $q^{(1)} \geq q$ and subsequent use-phase images with $q^{(2)} = q$, per the discussion in section "Rank-Ordered Detection Error Tradeoff Characteristics". The boxplot notch shows the median; the box shows the interquartile range, and the whiskers show the extreme values. Notches in (**d**) are not visible because the medians of FMRs are zero therefore outside the plot range.

the DETs for each quality level can be compared. Although a good QMA will exhibit an ordered relationship between quality and error rates, this DET computation is not operationally representative because an application cannot usually accept only samples with one quality value. Rather, the DET may be computed for verification of samples of quality k with enrollment samples of quality greater than or equal to k,

$$\mathrm{FNMR}(s,k) = \frac{\left|\left\{s_{ii} :\ s_{ii} \leq s,\ q_i^{(1)} \geq k, q_i^{(2)} = k\right\}\right|}{\left|\left\{s_{ii} :\ s_{ii} < \infty, q_i^{(1)} \geq k, q_i^{(2)} = k\right\}\right|},$$

$$\mathrm{FMR}(s,k) = \frac{\left|\left\{s_{ij} :\ s_{ij} > s,\ \ q_i^{(1)} \geq k, q_j^{(2)} = k, i \neq j\right\}\right|}{\left|\left\{s_{ij} :\ s_{ij} > -\infty,\ q_i^{(1)} \geq k, q_j^{(2)} = k, i \neq j\right\}\right|}, \quad (8)$$

The situation is modeled in which the enrollment samples are at least as good as the authentication (i.e., user submitted) samples. Such a use of quality would lead to ▶ failures to acquire for the low quality levels.

If instead performance across *all* authentication samples is compared against enrollment samples of quality greater than or equal to k,

$$\mathrm{FNMR}(s,k) = \frac{\left|\left\{s_{ii} :\ \ s_{ii} \leq s,\ \ q_i^{(1)} \geq k\right\}\right|}{\left|\left\{s_{ii} :\ \ s_{ii} < \infty,\ q_i^{(1)} \geq k\right\}\right|},$$

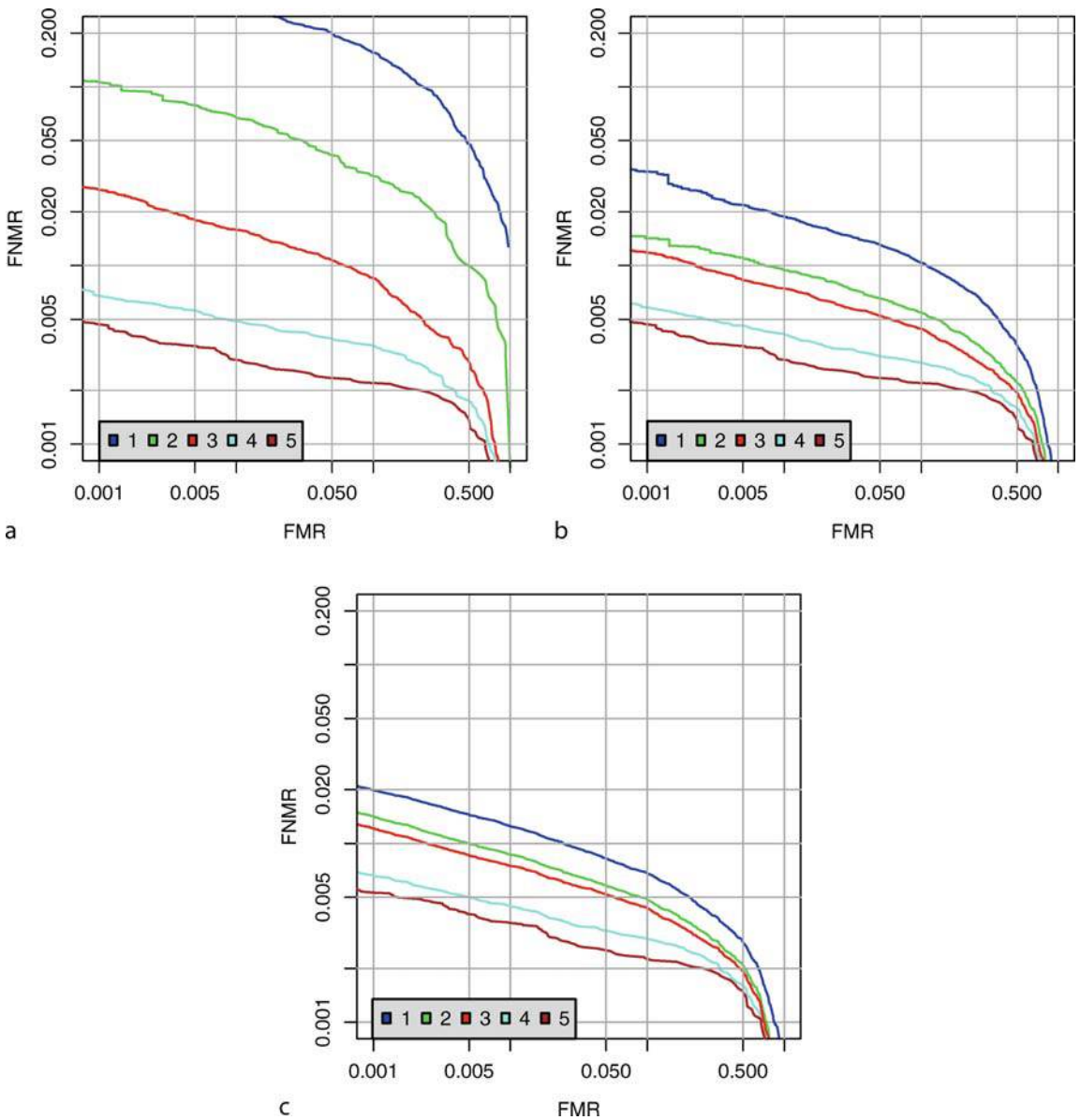
$$\mathrm{FMR}(s,k) = \frac{\left|\left\{s_{ij} :\ \ s_{ij} > s,\ \ \ q_i^{(1)} \geq k,\ i \neq j\right\}\right|}{\left|\left\{s_{ij} :\ \ s_{ij} > -\infty,\ \ q_i^{(1)} \geq k,\ i \neq j\right\}\right|}, \quad (9)$$

The situation where quality control is applied only during enrollment is modeled. If repeated enrollment attempts fail to produce a sample with quality above some threshold, a failure-to-enroll (FTE) would be declared. This scenario is common and possible

because enrollment, as an attended activity, tends to produce samples of better quality than authentication.

The considerable differences between these three formulations are evident in the DETs of Fig. 3 for which the NFIQ algorithm [6] for the predicting performance of a commercial fingerprint system was applied to over 61,993 genuine and 121,997 impostor comparisons (NFIQ native scores were transformed to $Q = 6 - \text{NFIQ}$). In all cases, the ranked separation of the DETs is excellent across all operating points. It is recommended that (8), as shown in Fig. 3(b), be used because it is a more realistic operational model.

However, as relevant as DET curves are to expected performance, revisited here is a very important



Biometric Sample Quality. Figure 3 Quality ranked detection error tradeoff characteristics. Each plot shows five traces corresponding to five transformed NFIQ levels. (Note that the DET used here plots FNMR vs. FMR on log scales. It is unconventional in that it does not transform the data by the CDF of the standard normal distribution. The receiver operating characteristic plots 1 − FNMR on a linear scale instead. These characteristics are used ubiquitously to summarize verification performance).

complication. Because DET characteristics quantify the separation of the genuine and impostor distributions and combine the effect of quality on both genuine and impostor performance, the separate effects of quality on FNMR and FMR is lost sight of.

In any case, it is concluded that DETs, while familiar and highly relevant, confound genuine and impostor scores. The alternative is to look at the specific dependence of the error rates on quality at some fixed threshold. Indeed for verification applications, the variation in FNMR with quality is key because the majority of transactions are genuine attempts. For negative identification systems (e.g., watchlist applications) in which users are usually not enrolled, the variation of FMR with quality is critical. This approach is followed in the next section.

Error Versus Reject Curves

It is proposed to use error versus reject curves as an alternative means of evaluating QMAs. The goal is to state how efficiently rejection of low quality samples results in improved performance. This again models the operational case in which quality is maintained by reacquisition after a low quality sample is detected. Consider that a pair of samples (from the same subject), with qualities $q_i^{(1)}$ and $q_i^{(2)}$, are compared to produce a score $s_{ii}^{(k)}$, and this is repeated for N such pairs.

Thresholds u and v are introduced that define levels of acceptable quality and define the set of low quality entries as

$$R(u,v) = \left\{ j \ : \ q_j^{(1)} < \ u, \quad q_j^{(2)} < v \right\}. \tag{10}$$

The FNMR is the fraction of genuine scores below threshold computed for those samples *not* in this set

$$\mathrm{FNMR}\,(t,u,v) = \frac{\left|\left\{ s_{jj} : s_{jj} \le t, j \notin R(u,v) \right\}\right|}{\left|\left\{ s_{jj} : s_{jj} < \infty \right\}\right|}. \tag{11}$$

The value of t is fixed (Note that any threshold may be used. Practically it will be set to give some reasonable false non-match rate, f, by using the quantile function the empirical cumulative distribution function of the genuine scores, $t = M^{-1}(1-f)$.) and u and v are varied to show the dependence of FNMR on quality.

For the one-dimensional case, when only one quality value is used the rejection set is

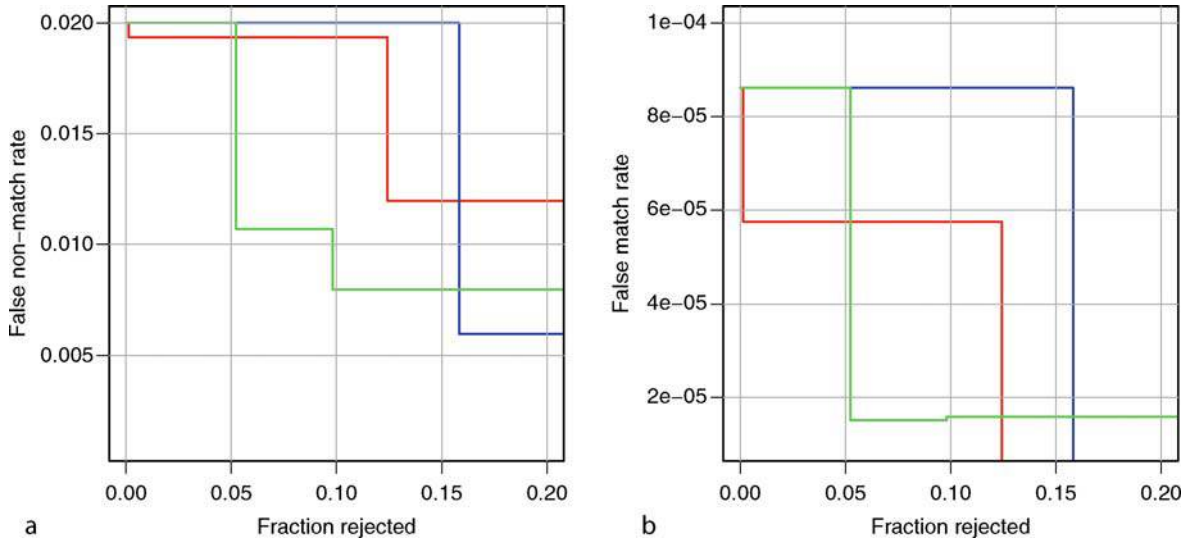
$$R(u) = \left\{ j \ : \ H(q_j^{(1)}, q_j^{(2)}) < u \right\} \tag{12}$$

where H is a function of combining two quality measures into a single measure. FNMR is false non-match performance as the proportion of nonexcluded scores below the threshold.

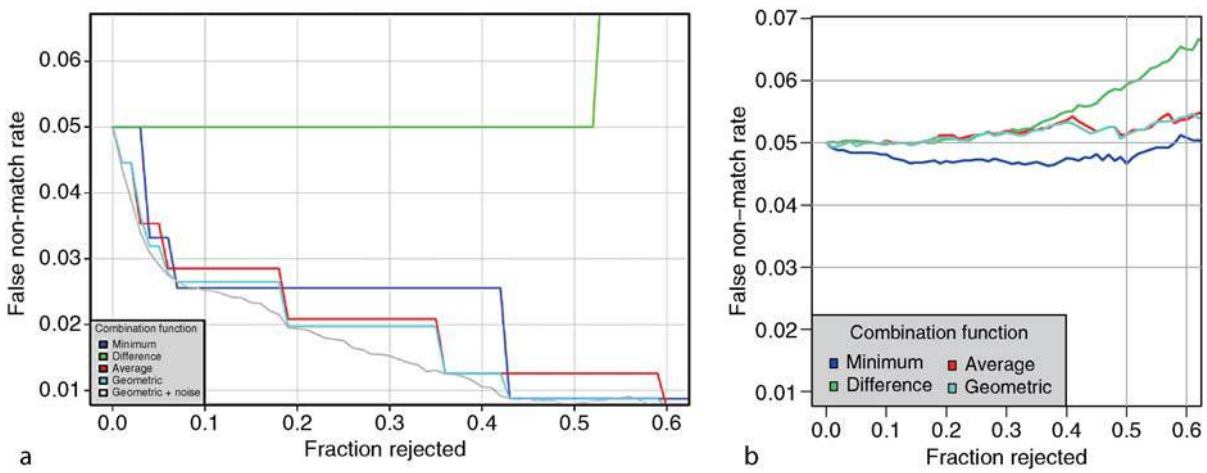
$$\mathrm{FNMR}(t,u) = \frac{\left|\left\{ s_{jj} : s_{jj} \le t, j \notin R(u) \right\}\right|}{\left|\left\{ s_{jj} : s_{jj} < \infty \right\}\right|} \tag{13}$$

If the quality values are perfectly correlated with the genuine scores, then when t is set to give an overall FNMR of x and then reject proportion x with the lowest qualities, a recomputation of FNMR should be zero. Thus, a good quality metric correctly labels those samples that cause low genuine scores as poor quality. For a good quality algorithm, FNMR should decrease quickly with the fraction rejected. The results of applying this analysis are shown in Fig. 4. Note that the curves for each of the three fingerprint quality algorithms trend in the correct direction, but that the even after rejection of 20% the FNMR value has fallen only by about a half from its starting point. Rejection of 20% is probably not an operational possibility unless an immediate reacquisition can yield better quality values for those persons. Yoshida, using the same approach, reported similar figures [18]. Note, however, that for NFIQ, the improvement is achieved after rejection of just 5%. In verification applications such as access control, the prior probability of an impostor transaction is low and thus, the overall error rate is governed by false nonmatchers. In such circumstances, correct detection of samples likely to be falsely rejected should drive the design of QMAs.

Figure 5 shows error versus reject behavior for the NFIQ quality method when the various $H(q_1, q_2)$ combination functions are used. Between the minimum, mean, and geometric mean functions there is little difference. The geometric mean is best (absent a significance test) with steps occurring at values corresponding to the square roots of the product of NFIQ values. The gray line in the figure shows $H = \sqrt{q_1 q_2} + N(0, 0.01)$, where the gaussian noise serves to randomly reject samples within a quality level and produces an approximation of the lower convex hull of the geometric mean curve. The green line result, for $H = \mid q_1 - q_2 \mid$, shows that transformed genuine comparison score is unrelated to the difference in the qualities of the samples. Instead, the conclusion is that FNMR is related to monotonic functions of the two values. The applicability of this result to other quality methods is not known.



Biometric Sample Quality. **Figure 4** Error versus reject performance for three fingerprint quality methods. (**a**) and (**b**) show reduction in FNMR and FMR at a fixed threshold as up to 20% of the low quality samples are rejected. The similarity scores come from a commercial matcher.



Biometric Sample Quality. **Figure 5** Dependence of the error versus reject characteristic on the quality combination function H(.). The plots show, for a fixed threshold, the decrease in FNMR as up to 60% of the low quality values are rejected. The similarity scores come from commercial matchers. The steps in (**a**) are result of discrete quality metric. Continuous quality metrics such as in (**b**), do not usually exhibit such steps.

Generalization to Multiple Matchers

It is a common contention that the efficacy of a quality algorithm is necessarily tied to a particular matcher. It is observed that this one-matcher case is commonplace and useful in a limited fashion and should, therefore, be subject to evaluation. However, it is also observed that it is possible for a quality algorithm to be capable of generalizing across *all* (or a class of) matchers, and this too should be evaluated.

Generality to multiple matchers can be thought of as an interoperability issue: can supplier A's quality measure be used with supplier B's matcher? Such a capability will exist to the extent that pathological

samples do present problems to both A and B's matching algorithms. However, the desirable property of generality exposes another problem: it cannot be expected to predict performance absolutely because there are good and bad matching systems. A system here includes all of the needed image analysis and comparison tasks. Rather, it is asserted that a quality algorithm intended to predict performance generally need only be capable of giving a relative or rank ordering, i.e., low quality samples should give lower performance than high quality samples.

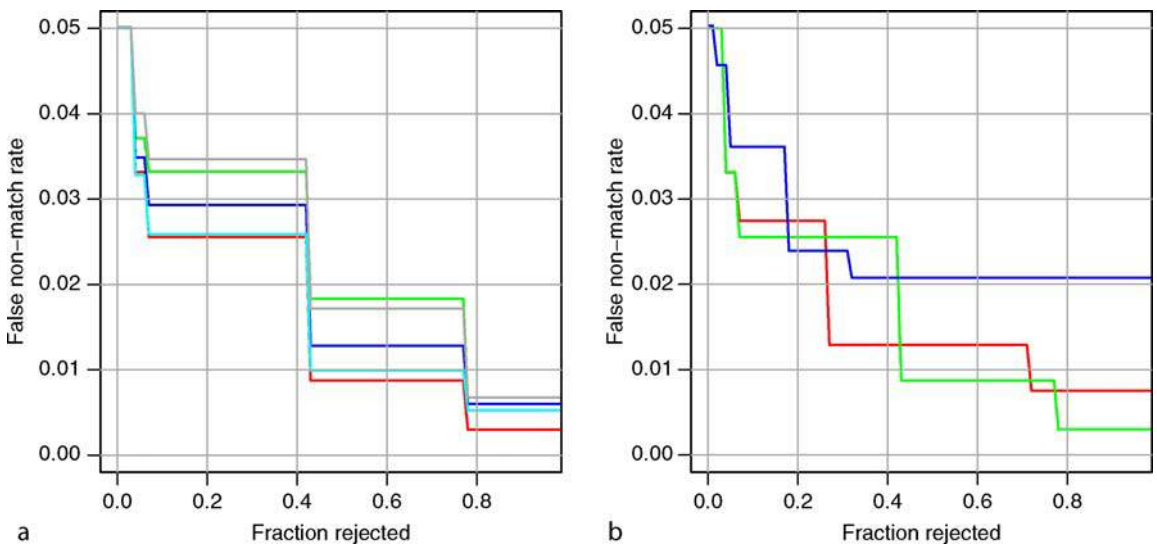
The plots of Fig. 6 quantify this generalization for the NFIQ algorithm using the error versus reject curves of section "Error Versus Reject Curves". Figure 6(a) includes five traces, one for each of five verification algorithms. The vertical spread of the traces indicates some disparity in how well NFIQ predicts the performance of the five matchers. A perfectly general QMA would produce no spread.

Measuring Separation of Genuine and Impostor Distributions

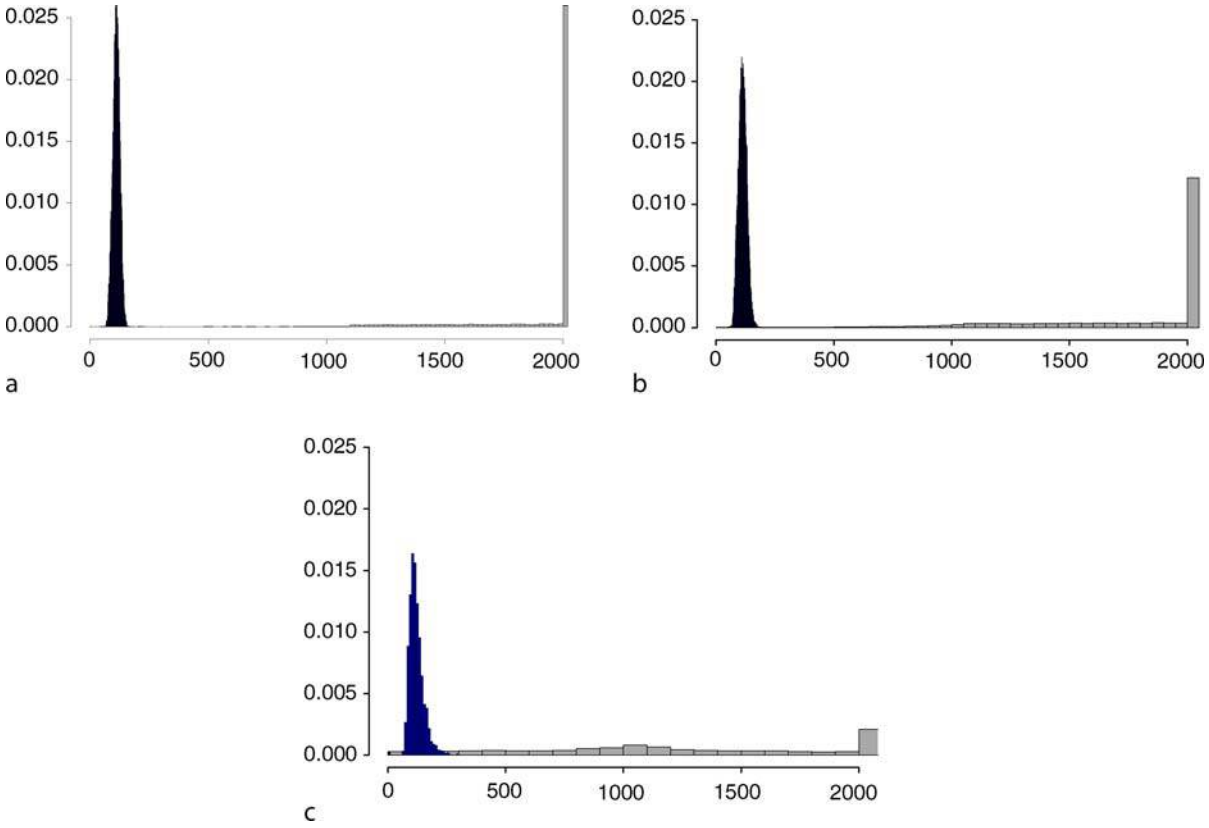
Quality algorithms can be evaluated on their ability to predict how far a genuine score will lie from its impostor distribution. This means instead of evaluating a quality algorithm solely based on its FNMR (i.e., genuine score distribution), the evaluation can be augmented by including a measure of FMR because correct identification of an enrolled user depends both on correctly finding the match and on rejecting the nonmatches. Note also that a quality algorithm could invoke a matcher to compare the input sample with some internal background samples to compute sample mean and standard deviation.

The plots of Fig. 7 show, respectively, the genuine and impostor distributions for adjusted NFIQ values, 1, 3, and 5. The overlapping of genuine and impostor distributions for the poorest NFIQ means higher recognition errors for that NFIQ level, and vice versa; the almost complete separation of the two distribution for the best quality samples indicates lower recognition error. NFIQ was trained to specifically exhibit this behavior.

The Kolmogorov–Smirnov is considered statistic. For better quality samples, a larger KS test statistic (i.e., higher separation between genuine and impostor distribution) is expected. Each row of Table 1 shows KS statistics for one of the three quality algorithms tested. KS statistics for each quality levels $u = 1, \ldots, 5$ are computed by first computing the genuine (i.e., $\{s_{ii}: (i, i) \in R(u)\}$) and impostor (i.e., $\{s_{ij}: (i, j) \in R(u), i \neq j\}$) empirical cumulative distributions, where $R(u) = \{(i,j): H(q_i^{(1)}, q_j^{(2)}) = u\}$. Thereafter, the largest absolute difference between the genuine and impostor



Biometric Sample Quality. Figure 6 Error versus reject characteristics showing how NFIQ generalizes across (**a**) five verification algorithms and (**b**) three operational data sets. The steps in (**a**) occur at the same rejection values because the matchers were run on a common database.



Biometric Sample Quality. Figure 7 There is a higher degree of separation between the genuine and impostor distribution for better quality samples as measured by NFIQ.

Biometric Sample Quality. Table 1 KS statistics for quality levels of three quality algorithms

Quality algorithm	$Q = 1$	$Q = 2$	$Q = 3$	$Q = 4$
Quality algorithm 1	0.649	0.970	0.988	0.993
Quality algorithm 2	0.959	0.995	0.996	0.997
Quality algorithm 3	0.918	0.981	0.994	0.997

distributions of quality u is measured and plotted. (Note that to keep quality algorithm providers anonymous KS statistics of the lowest four quality levels were reported.)

Summary

Biometric quality measurement is an operationally important and difficult problem that is nevertheless massively under-researched, in comparison to the primary feature extraction and pattern recognition tasks. It was asserted that quality algorithms should be developed to explicitly target matching error rates, and not human perceptions of sample quality.

Several means were given for assessing the efficacy of quality algorithms. The existing practice was reviewed, cautioned against the use of detection error tradeoff characteristics as the primary metrics, and instead advanced boxplots and error versus reject curves as preferable. This chapter suggests that algorithm designers should target false non-match rate as the primary performance indicator.

Related Entries

- ▶ Authentication
- ▶ Biometric Sample Quality, Standardization
- ▶ Biometric Systems, Agent-Based
- ▶ Biometric Vulnerabilities, Overview
- ▶ Biometric Quality
- ▶ Enrollment

- ▶ Face Image Quality
- ▶ Fingerprint Image Quality
- ▶ Identification
- ▶ Iris Image Quality
- ▶ Verification

References

1. Mansfield, A.J.: ISO/IEC 19795-1 Biometric Performance Testing and Reporting: Principles and Framework, FDIS ed., JTC1/SC37/Working Group 5, August 2005, http://isotc.iso.org/isotcportal
2. Ko, T., Krishnan, R.: Monitoring and reporting of fingerprint image quality and match accuracy for a large user application. In: Proceedings of the 33rd Applied Image Pattern Recognition Workshop. IEEE Computer Society, pp. 159–164 (2004)
3. Proceedings of the NIST Biometric Quality Workshop. NIST (March 2006), http://www.itl.nist.gov/iad/894.03/quality/workshop/presentations.html
4. Benini, D., et al.: ISO/IEC 29794-1 Biometric Quality Framework Standard, 1st ed. JTC1/SC37/Working Group 3 (Jan 2006), http://isotc.iso.org/isotcportal
5. Chen, Y., Dass, S., Jain, A.: Fingerprint quality indices for predicting authentication performance. In: Procedings of the Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 160–170 (July 2005)
6. Tabassi, E.: Fingerprint Image Quality, NFIQ, NISTIR 7151 ed., National Institute of Standards and Technology (2004)
7. Bioscrypt Inc., Systems and Methods with Identify Verification by Comparison and Interpretation of Skin Patterns Such as Fingerprints, http://www.bioscrypt.com (June 1999)
8. Alonso-Fernandez, F., Fierrez-Aguilar, J., Ortega-Garcia, J., A review of schemes for fingerprint image quality computation. In COST 275 – Biometrics-based recognition of people over the internet (October 2005)
9. Lim, E., Jiang, X., Yau, W.: Fingerprint quality and validity analysis. In: Proceedings of the IEEE Conference on Image Processing, vol. 1, pp. 469–472 (September 2002)
10. Tilton, C., et al.: The BioAPI Specification, American National Standards Institute, Inc. (2002)
11. ISO/IEC JTC1/SC37/Working Group 3, ISO/IEC 19794 Biometric Data Interchange Formats, http://isotc.iso.org/isotcportal (2005)
12. Tabassi, E.: A novel approach to fingerprint image quality. In: IEEE International Conference on Image Processing ICIP-05, Genoa, Italy (September 2005)
13. Chambers, J.M., Cleveland, W.S., Kleiner, B., Tukey, P.A.: Graphical Methods for Data Analysis, p. 62. Wadsworth and Brooks/Cole (1983)
14. Fierrez-Aguilar, J., Muñoz-Serrano, L., Alonso-Fernandez, F., Ortega-Garcia, J.: On the effects of image quality degradation on minutiae and ridge-based automatic fingerprint recognition. In IEEE International Carnahan Conference on Security Technology (October 2005)
15. Martin, A., Doddington, G.R., Kamm, T., Ordowski, M., Przybocki, M.A.: The DET curve in assessment of detection task performance. In: Proceedings of Eurospeech, pp. 1895–1898. Rhodes, Italy, Greece (1997)
16. Mansfield, A.J., Wayman, J.L.: Best practices in testing and reporting performance of biometric devices. National Physics Laboratory Report CMSC 14/02, August 2002, http://www.cesg.gov.uk/site/ast/biometrics/media/BestPractice.pdf (2002)
17. Simon-Zorita, D., Ortega-Garcia, J., Fierrez-Aguilar, J., Gonzalez-Rodriguez, J.: Image quality and position variability assessment in minutiae-based fingerprint verification. IEE Proceedings on Vision, Image and Signal Processing, vol. 150, no. 6, pp. 395–401, December 2003, special Issue on Biometrics on the Internet (2003)
18. Yoshida, A., Hara, M.: Fingerprint image quality metrics that guarantees matching accuracy. In: Procedings of NIST Biometric Quality Workshop. NEC Corp., March 2006, http://www.itl.nist.gov/iad/894.03/quality/workshop/presentations.html

Biometric Sample Quality, Standardization

Elham Tabassi, Patrick Grother
National Institute of Standards and Technology, MD, USA

Synonyms

Biometric quality; Sample quality

Definition

Open documented data structures for universally interpretable interchange of ▶ biometric sample quality data.

▶ Biometric data interchange standards are needed to allow the recipient of a data record to successfully process data from an arbitrary producer. This defines biometric interoperability and the connotation of the phrase "successfully process" the data, in this case, ▶ Biometric Sample Quality score, can be accurately exchanged and interpreted by different applications. This can only be achieved if the data record is both syntactically and semantically conformant to the documentary standard.

Introduction

Performance of biometric systems depends on the quality of the acquired input samples. If quality can be improved, either by sensor design, user interface design, or by standards compliance, better performance can be realized. For those aspects of quality that cannot be designed-in, an ability to analyze the quality of a live sample is needed. This is useful primarily in initiating the reacquisition from a user, but also for the real-time selection of the best sample, and the selective invocation of different processing methods. That is why quality measurement algorithms are increasingly deployed in operational biometric systems. With the increase in deployment of quality algorithms, rises the need to standardize an interoperable way to store and exchange of biometric quality scores.

Roles

With advancement in biometric technologies as a reliable identity authentication scheme, more large-scale deployments (e.g., e-passport) involving multiple organizations and suppliers are being ruled out. Therefore, in response to a need for interoperability, biometric standards have been developed.

Without interoperable biometric data standards, exchange of biometric data among different applications is not possible. Seamless data sharing is essential to identity management applications when enrollment, capture, searching, and screening are done by different agencies, at different times, using different equipment in different environments and/or locations. Interoperability allows modular integration of products without compromising architectural scope, and facilitates the upgrade process and thereby mitigates against obsolescence.

This chapter focuses on biometric quality standardization. Broadly biometric quality standards serve the same purpose as many other standards, which is to establish an interoperable definition, interpretation, and exchange of biometric quality data. Like other standards, this creates grounds for a marketplace of off-the-shelf products, and is a necessary condition to achieve supplier independence, and to avoid vendor lock-in.

Biometric quality measurement has vital roles to play in improving biometric system accuracy and efficiency during the capture process (as a control-loop variable to initiate reacquisition), in database maintenance (sample update), in enterprise wide quality-assurance surveying, and in invocation of quality-directed processing of samples. Neglecting quality measurement will adversely impact accuracy and efficiency of biometric recognition systems (e.g., verification and identification of individuals). Accordingly, biometric quality measurement algorithms are increasingly deployed in operational systems [1, 2]. These motivated for biometric quality standardization efforts.

Standards do not themselves assure interoperability. Specifically, when a standard is not fully prescriptive, or allows for optional content, then two implementations that are exactly conformant to the standard may still not interoperate. This situation may be averted by applying further constraints on the application of the standard. This is done by means of "application profile" standards which formally call out the needed base standards and refine their optional content and interpretation.

Standards Development Organizations

Standards are developed by a multitude of standards development organizations (SDOs) operating in a great variety of technical disciplines. SDO's exist within companies and governments, and underneath trade associations and international body umbrellas. International standards promise to support larger marketplaces and the development process involves more diverse and thorough review and so consensus is more difficult to achieve. Standard development processes are conducted according to definitive sets of rules. These are intended to achieve consensus standards that are technically sound, implementable, and effective.

The following list gives an overview of the relevant SDOs. Note that the published standards are usually copyrighted documents and available only by purchase.

- ISO/IEC JTC 1/SC 37: SubCommittee 37 (SC 37) *Biometrics* was established in mid 2002 as the most new of seventeen active subcommittees beneath Joint Technical Committee 1 (JTC 1) and its parent the International Standard Organization (ISO) and the International Electrotechnical Commission (IEC) (ISO maintains a catalog of its standards development efforts at http://www.iso.org/iso/en/CatalogueListPage.CatalogueList). The scope of

JTC 1/SC 37 is standardization of generic biometric technologies pertaining to human beings to support interoperability and data interchange among applications and systems. The establishment of JTC 1/SC 37 provided an international venue to accelerate and harmonize formal international biometric standardization and to coalesce a wide range of interests among information technology and biometric industry and users of biometric-based solutions for multiple identification and verification applications. SC 37 portfolio is divided into six working groups of SC 37. The body responsible for biometric quality standardization is Working Group 3. The group is the largest Working Group in SC 37 and develops biometric data interchange format standards, which have the highest profile adoption in the marketplace.

- M1: M1 is Technical Committee of the InterNational Committee for Information Technology Standards (INCITS). It serves as the United States Technical Advisory Group (TAG) to SC 37. It was established in June 2002 and is responsible for formulating U.S. positions in SC 37 where it holds the U.S. vote. It is also a standards development organization in its own right. Its standards are published in the US, but may be purchased worldwide.
- ANSI/NIST The U.S. National Institute of Standards and Technology (NIST) is also a SDO. It developed the ANSI/NIST standards for law enforcement under the canvass process defined by American National Standard Institution (ANSI).

The ISO/IEC 29794 Biometric Sample Quality Standard

In January 2006, the SC37 Biometrics Subcommittee of JTC1 initiated work on ISO/IEC 29794, a multipart standard that establishes quality requirements for generic aspects (Part 1), fingerprint image (Part 4), facial image (Part 5), and possibly, other biometrics later. Specifically, part 1 of this multi-part standard specifies derivation, expression, and interpretation of biometric quality regardless of modality. It also addresses the interchange of biometric quality data via the multipart ISO/IEC 19794 Biometric Data Interchange Format Standard [4]. Parts 4 and 5 are technical reports (not standard drafts) which address the aspects of biometric sample quality that are specific to finger images and facial images as defined in ISO/IEC 19794-4 and ISO/IEC 19794-5 respectively.

The generic ISO quality draft (ISO/IEC 29794-1) requires that quality values must be indicative of recognition performance in terms of false match rate, false non-match rate, failure to enrol and failure to acquire. Furthermore, it considers three components of biometric sample quality namely character, fidelity and utility. The character of a sample indicates the richness of features and traits from which the biometric sample is derived. The fidelity of a sample is defined as the degree of similarity between a biometric sample and its source, for example, a heavily compressed fingerprint has low fidelity. The utility of a sample reflects the observed or predicted positive or negative contribution of an individual sample to the overall performance of a biometric system. Utility is a function of both the character and fidelity of a sample and is most closely indicative of performance in terms of recognition error rates (i.e., false match rate, false non-match rate, failure to enrol and failure to acquire).

Part 1 of multipart ISO/IEC 29794 draft standard defines a binary record structure for the storage of a sample's quality data. It establishes requirements on the syntax and semantic content of the structure. Specifically it states that the purpose of assigning a quality score to a biometric sample shall be to indicate the expected utility of that sample in an automated comparison environment. That is, a quality algorithm should produce quality scores that target application specific performance variables. For verification, the metric would usually be false-match and false-non-match rates that are likely to be realized when the sample is matched.

In addition, revision of all parts of ISO/IEC 19794 Biometric Data Interchange Format began in January 2007. This opened the opportunity to revise or add quality-related clauses (e.g., compression limits) to data format standards so that conformance to those standards ensures acquisition of sufficient quality samples. This constitutes quality by-design. To enable an interoperable way of reporting and exchanging biometric data quality scores, the inclusion of a five-byte quality field to the view header in each view of the data in a Biometric Data Block (BDB) for all parts of ISO/IEC 19794 is being considered. By placing quality field in the view header (as opposed to general header) of a BDB, one can precisely report quality score for each view of a biometric sample (Fig. 1). Table 1 shows the

Biometric Sample Quality, Standardization. **Figure 1** Structure of header in a biometric data block as defined in ISO/IEC 19794-x.

Biometric Sample Quality, Standardization. **Table 1** Structure of five-byte quality field that SC 37 Working Group 3 is considering

Description	Size (byte)	Valid values	Note
Quality Score	1	[0-100] 255	0: lowest; 100: highest; 255: Failed Attempt
Quality Algorithm Vendor ID	2	[1,65535]	These two bytes uniquely identifies the supplier (vendor) of quality score
Quality Algorithm ID	2	[1,65535]	These two bytes uniquely identifies the algorithm that computes the quality score. It is provided by the supplier (vendor) of quality score

structure of the quality filed that SC 37 Working Group 3 is currently considering.

The one-byte quality score shall be a quantitative expression of the predicted matching performance of the biometric sample. Valid values for quality score are integers between 0 and 100, where higher values indicate better quality. Value 255 is to handle special cases. An entry of "255" shall indicate a failed attempt to calculate a quality score. This values of quality score is harmonized with ISO/IEC 19784-1 BioAPI Specification (section 0.5) [6], where "255" is equivalent to BioAPI "-1" (Note that BioAPI, unlike ISO/IEC 19794 uses signed integers).

To enable the recipient of the quality score to differentiate between quality scores generated by different algorithms, the provider of quality scores shall be uniquely identified by the two most significant bytes of four-byte Quality Algorithm vendor ID (QAID). The least significant two bytes shall specify an integer product code assigned by the vendor of the quality algorithm. It indicates which of the vendors algorithms (and version) was used in the calculation of the quality score and should be within the range 1 – 65535.

Different quality assessment methods could be used to assess quality of a biometric sample, for example, quality algorithm A could be used at the time of enrollment, but the verification phase might deploy quality algorithm B. To accomodate interchange of quality scores computed by different quality algirithms, multiple blocks of quality as shown in Table 1 could be encoded in a view header. Block(s) of quality data as shown in Table 1 is preceded by a single byte which value indicates how many blocks of quality data are to follow. A value of O means no attempt was made to calculate a quality score (i.e. no quality score has been specified). This is equivalent to BioAPI "−2". The structure of the quality field is modality independent and therefore generalizable to all parts of ISO/IEC 19794.

The ISO/IEC 29794 standard is currently under development, and ISO/IEC 19794 is currently under revision. The reader is cautioned that standards under development or revision, are subject to change; the documents are owned by the respective working groups and their content can shift due to various reasons including, but not limited to technical difficulties, the level of support, or the need to gain consensus.

The ANSI/NIST ITL 1-2007 Quality Field

Initiated in 1986, this standard is the earliest and most widely deployed biometric standard. It establishes formats for the markup and transmission of textual, minutia, and image data between law enforcement agencies, both within United States and internationally.

Biometric Sample Quality, Standardization. Table 2 BioAPI quality categories

Value	Interpretation
0-25	*Unacceptable*: The sample cannot be used for the purpose specified by the application. The sample needs to be replaced using one or more new biometric samples.
26-50	*Marginal*: The sample will provide poor performance for the purpose specified by the application and in most application environments will compromise the intent of the application. The sample needs to be replaced using one or more new biometric samples.
51-75	*Adequate*: The biometric data will provide good performance in most application environments based on the purpose specified by the application. The application should attempt to obtain higher quality data if the application developer anticipates demanding usage.
76-100	*Excellent*: The biometric data will provide good performance for the purpose specified by the application.

The ANSI/NIST standard includes defined *Types* for the major biometric modalities. The standard is multimodal in that it allows a user to define a transaction that would require, for example, fingerprint data as Type 14, a facial mugshot as Type 10, and the mandatory header and metadata records Type 1 and 2. These are linked with a common numeric identifier.

In its latest revision [8], the standard adopted the ISO five-byte quality field (Table 1) structure, but unlike ISO/IEC 29794, it allows for multiple quality fields, where each quality score could be computed by a different quality algorithm supplier. In addition, it mandates NIST Fingerprint Image Quality (NFIQ) [9] for all Type 14 records.

The BioAPI Quality Specification

ISO/IEC 19784 Biometric Application Programming Interface (BioAPI) [7] (and its national counterpart The BipAPI specification [6]) allows for quality measurements as an integral value in the range of 0–100 with exceptions that value of "-1" means that the quality field was not set by the Biometric Service Provider (BSP) and value of "-2" means that quality information is not supported by the BSP. The primary objective of quality measurement and reporting is to have the BSP inform the application how suitable the biometric sample is for the purpose specified by the application (as intended by the BSP implementer based on the use scenario envisioned by that BSP implementer), and the secondary objective is to provide the application with relative results (e.g., current sample is better/worse than previous sample). BioAPI also provides guidance on general interpretation of quality scores as shown in Table 2.

Summary

The benefit of measuring and reporting of biometric sample quality is to improve performance of biometric systems by improving the integrity of biometric databases and enabling quality-directed processing in particular when utilizing multiple biometrics. Such processing enhancements result in increasing probability of detection and track accuracy while decreasing probability of false alarms. Given these important roles of biometric sample quality in improving accuracy and efficiency of biometric systems, quality measurement algorithms are increasingly deployed in operational systems. Biometric Sample Quality standards have been developed to facilitate universal seemless exchange of sample quality information.

Related Entries

- ▶ Face Image Quality Assessment Software
- ▶ Face Sample Quality
- ▶ Fingerprint Image Quality
- ▶ Fusion, Quality-Based
- ▶ Interoperability
- ▶ Iris Image Quality

References

1. T. Ko, T., Krishnan, R.: Monitoring and reporting of fingerprint image quality and match accuracy for a large user application. In: Proceedings of the 33-rd Applied Image Pattern Recognition Workshop. IEEE Computer Society, pp. 159–164 (2004)
2. B. Scott Swann, Integrating Standard Biometric Quality Metric within the FBI IAFIS, In Proceedings of the NIST Biometric Quality Workshop. NIST, http://www.itl.nist.gov/

iad/894.03/quality/workshop/presentations.html, March 2006. Bradford Wing, Why is Biometric Quality Important to DNS and other Government Agencies, In Proceedings of the NIST Biometric Quality Workshop. NIST, http://www.itl.nist.gov/lad/894.03/quality/workshop/presentations.html, March 2006.
3. Procedings of the NIST Biometric Quality Workshop. NIST, Mar 2006, http://www.itl.nist.gov/iad/894.03/quality/workshop/presentations.html
4. ISO/IEC JTC1/SC37/Working Group 3: ISO/IEC 19794 Biometric Data Interchange Formats (2005)
5. Tilton, C., et al.: The BioAPI Specification, American National Standards Institute, Inc (2002)
6. ISO/IEC JTC1/SC37/Working Group 3: ISO/IEC 19784-1 Biometric application programming interface with Amd. 1 (2008)
7. McCabe, R.M., et al.: Data Format for the interchange of Fingerprint, Facial, and Other Biometric Information, ANSI/NIST (2007)
8. Tabassi, E., et al.: Fingerprint Image Quality, NFIQ, NISTIR 7151 ed., National Institute of Standards and Technology (2004)

Biometric Sample Synthesis

Douglas J. Buettner
The Aerospace Corporation, El Segundo, CA, USA

Synonyms

Synthetic biometrics; Artificial biometrics; Artificial digital biometrics; Artificial image biometrics; Intermediate biometrics

Definition

Biometric sample synthesis is the computer generation of simulated digital biometric data using parametric models. Parametric models are in general the computer creation steps derived from the empirical analysis of digitized biometric patterns or mathematical equations from the physics of the biometric sample's creation.

Introduction

Biometric sample synthesis is the art and science of creating artificial digital biometrics that mimic real digital biometric samples. Researchers involved in the creation of synthetic biometric samples may have any number of possible noble goals; included in these may be striving for a fundamental understanding of the factors that affect the digitization process of real human biometric samples for a specific type of biometric sensor, attempting to improve or test computer algorithms used in biometric security devices, striving for statistically realistic equations of human populations, or simply attempting to efficiently computer-generate an image that is similar in visual appearance to a digitized biometric image.

No matter what the underlying reason for creating synthetic biometrics by researchers, the movie industry's quest for realistic computer-generated artificial personas has led to ▶ physics-based models to control physical form, motion, and illumination properties of materials [1]. Computer-generated human characteristics now address a broad range of human details including facial features, skin, hair, and gait, as well as more nuanced bodily movements, such as emotive gestures and even eye movement. The Association for Computing Machinery (ACM) Special Interest Group on Computer Graphics (SIGGRAPH) has a large body of work spanning over three decades with the long-standing goal of achieving photo-realism in the computer generation of synthetic images [2]. This achievement of modeling, animation, and rendering of visual human subjects is widely viewed in feature films, commercial art, and video games. An example of the state-of-the-art in the synthesis of an image-based facial biometrics is illustrated in Fig. 1.

The ultimate goal of biometric sample synthesis can be summarized as; the use of a standard computer model containing parameter settings that provide the ability to create a synthetic corpus of biometrics, which would be indistinguishable from that of a corpus of biometric samples obtained from real people.

Factors Affecting Biometric Samples

There are a number of factors that directly affect real biometric samples, which the process of biometric sample synthesis must take into account. For example, biological human responses to environmental conditions are known to directly influence a biometric sample such as: heat to sweat, cold to shivering, or light level on pupil dilation. Likewise, the environment can

Biometric Sample Synthesis. **Figure 1** Rendering of a synthetic face using 13 million triangles and a bidirectional surface scattering distribution function (BSSRDF) model for subsurface light scattering and an oily reflection layer (http://graphics.ucsd.edu/~henrik/papers/face_cloning/) (*Reproduced with permission from the author).*

also directly affect the biometric digitizing device; for example, fog, rain, smoke or light level decrease a video camera's ability to get a clear image, or water on a fingerprint device's platen can adversely affect the quality of the image. The environment can also cause behavioral changes that affect biometric sample acquisition; for example, influencing the clothes we wear during hot or cold weather or during a cloudy or sunny day, or whether or not we are likely to be wearing sunglasses, gloves, or certain types of headgear. Additionally, one's occupation can affect the exposure of a biometric to specific environments that may degrade the quality of the biometric sample. The impact of handling rough surfaces on the skin ridges and troughs on the fingers and hands of people in certain occupations can directly affect the quality of biometric samples from some biometric fingerprint digitizing devices.

Regional location also affects the likelihood of finding various ethnicities who may wear different kinds of hats, different styles of facial hair growth, or various types of garments, which may directly influence biometric sample acquisition. Additional factors may effect biometric sample acquisition through the presentation of a biometric to the digitizing device. An example is the habituation of users to fingerprint sensor technologies that require pressing the sensor's platen; users unfamiliar with the technology are more likely to press extremely hard or very lightly, while habituated users are more likely to provide a closer to nominal amount of pressure when placing a digit on the device. The amount of pressure may (or may not) adversely affect the biometric feature extraction algorithm used by the vendor. For example, light pressure could decrease the number of minutia available to the biometric matching algorithm.

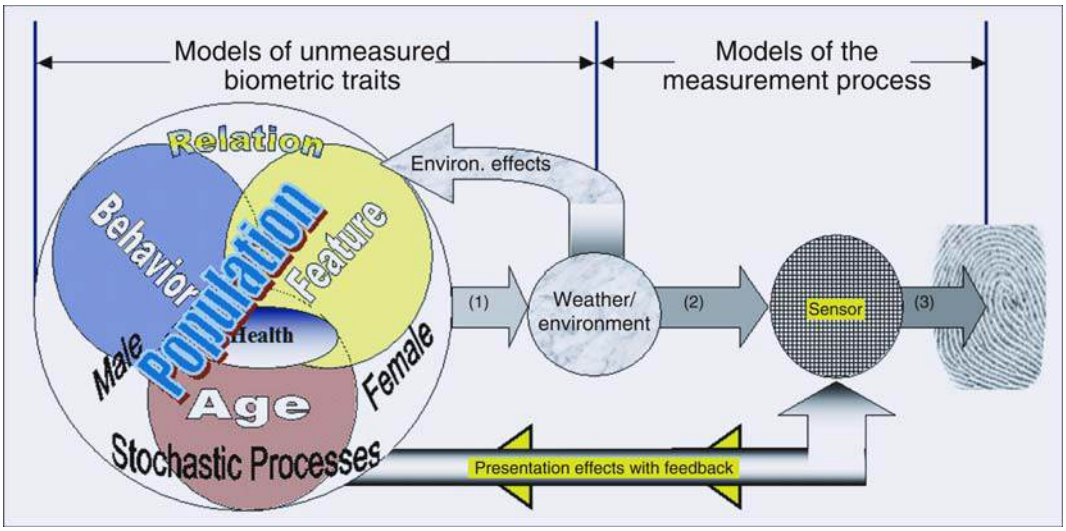
Genetic factors also play an important role in biometrics. Examples here include the generally smaller size of Asian fingerprints, gender, skin color, and others. Another environmental factor that can affect a biometric is our health, in the sense that our unique genetic makeup and our environmental exposure to triggering factors can make us more susceptible to (for example) diseases that can affect a biometric. Here an example is psoriasis that can affect the skin, which (if located on the hand or fingers) can affect the quality of a finger or palm acquisition device's digitized image that in turn can affect the ability of the biometric feature extraction algorithm to extract a consistent biometric feature. Finally, the natural process of aging and relationships with exposure to the sun affects the quality or number of features available for biometric matching algorithms.

The method of sample measurement also directly affects the quality and depth of information obtained about the real biometric trait that the device is attempting to measure. Examples are optical, electrical resistance, or ultrasound for fingerprint devices, and number of pixels used by a digital-camera to acquire images of the face. The final representation of the synthetic biometric sample must adequately mimic the digitization process on the biometric sample. Figure 2 illustrates the taxonomy framework that distinguishes between the feedback effects of environment on unmeasured biometric samples and the measurement/digitization process [3].

How all these factors directly or indirectly affect biometric samples is an ongoing research activity in the field of biometrics.

Synthesis Methods

Synthesis of image-based biometrics has been achieved for the most widely recognized digital-image type



Biometric Sample Synthesis. **Figure 2** A conceptual biometric-environment-sensor interaction model for understanding the taxonomy of modeled parameters in synthetic biometrics.

biometrics of fingerprint, face, and iris. Table 1 identifies the available model types for a number of widely used biometric modalities.

The methods used for biometric sample synthesis can be categorized depending on the approach for feature synthesis. These are loosely placed into statistical or physical modeling categories based on characteristics of the biometric formation process.

Physical models are those that are based on the physics of how the biometric is created. Examples of biometric features that have physical models for the body part containing the biometric sample include stress/strain finger growth models that have been used to describe fingerprint patterns, craniofacial 3D growth models, and speech synthesis models for the human vocal tract.

► Statistical models are those that use ► empirical analysis of real 2D or 3D biometric images to create empirically derived statistical information that can be parameterized into some sort of equation or algorithmic synthesis steps to create a synthetic biometric sample. The SFinGe fingerprint generation tool in Fig. 3 is one example of the use of this intermediate-pattern type of biometric feature generator. This tool also exemplifies the parametric or mathematical model of synthesis. Face creation and morphing tools, such as the one from FaceGen Modeler from Singular Inversions, Inc. (Fig. 4) is another example of a statistical modeling tool that also provides age progression functionality as well as the ability to rotate, translate, add texture, or make a number of possible modifications to face/head models.

Validated statistical models would (at a minimum) be those models that have been rigorously validated to match across a wide range of human ethnic populations under specific image-gathering conditions that could affect the image. Matching could be achieved by using quantile-quantile (q-q) plots to show that the distributions from the two different populations are identical as was done by Daugman for iris codes [4]. The broader view would be the validation of these models across the human populations and the widest variety of possible environments and devices.

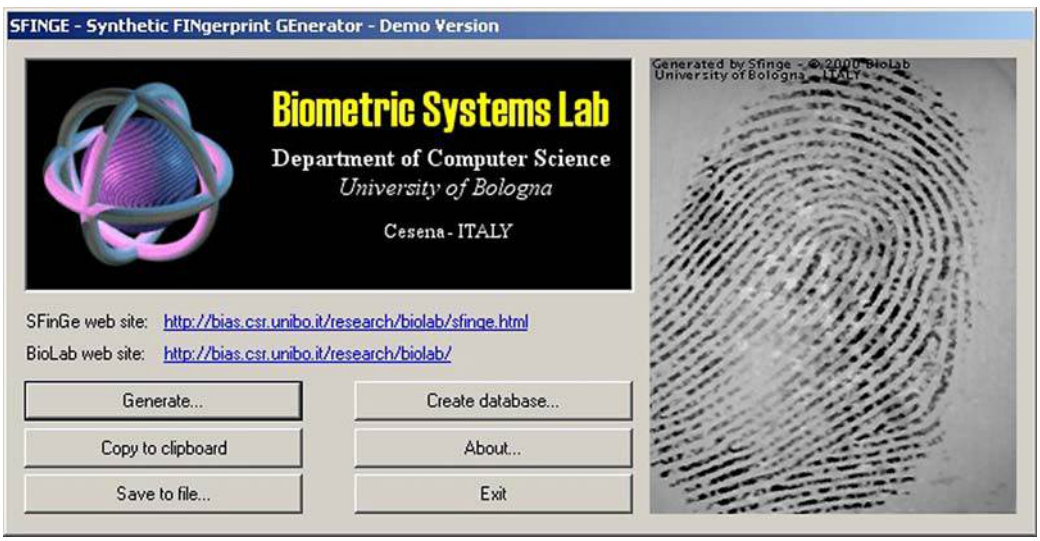
After the statistical taxonomy, parameters are understood for the acquisition of a particular biometric sample, the tool can be configured to generate a large number of synthetic biometric samples (as shown in Fig. 5).

Uses of Synthetic Biometrics

Synthetic biometric samples should not be considered a replacement for real biometric samples, which are still needed to understand the specifics of how the biometric acquisition device and system as a whole handles real world conditions. Mansfield and Wayman provide a warning about the "external validity" from the use of artificial images due to the bias that can result from their generation [5]. This bias is introduced

Biometric Sample Synthesis. **Table 1** Synthetic biometric data generation

	Fingerprint	Face	Iris	Voice
Synthetic generation	Yes	Yes	Yes	Yes
Model types	Physical – finger/skin growth model; Statistical – level 2 minutiae	Physical – craniofacial growth & human skin light scattering models; Statistical – morphable feature	Statistical – feature	Statistical and Physical – articulatory
Validated statistical models	No	No	Partial	No



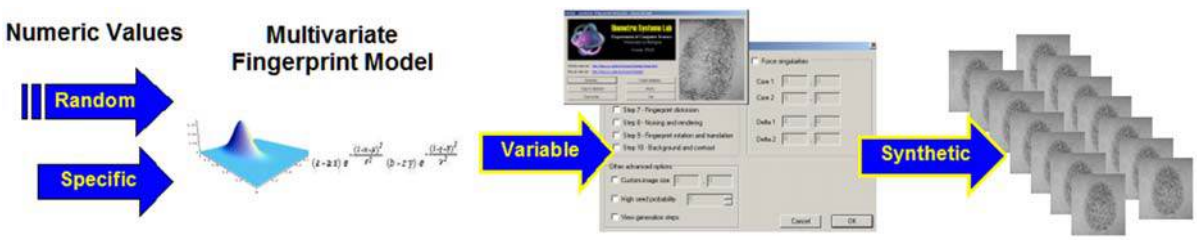
Biometric Sample Synthesis. **Figure 3** The SFinGe tool as an example of the synthesized intermediate biometric patterns based on an empirical statistical model.

during the analysis of the training set of images used in creating the parametric equations. However, synthetic biometric samples offer a number of potential uses, some positive and some negative.

Among the positive useful benefits of synthetic biometric sample is a cost-effective means for studying a biometric system's algorithmic sensitivity to specific biometric images from a variety of sensor types, or the performance impacts from biometric images that have been affected by any of a number of various environmental or presentation factors.

Injecting synthetic biometric images into real world or synthetic world scenes provides an ability to perform operational scenario testing in a laboratory environment. The modeled "subjects" can be generated randomly according to statistical models of a target population or in this randomly generated target population, very specific real or synthetic biometrics can be injected to determine gross failure to detect rates in a system context. Operational scenarios can include videos of synthetic subjects walking through a security checkpoint, and in the system context could include a specifically injected individual with specific behavioral characteristics, or individual wearing troublesome garments, such as sunglasses. Validated biometric models could also be used in the area of fingerprinting where they could readily provide the effects of age, ethnicity, and gender on performance. Biometric systems engineers could run a vast array of potential scenarios to categorize the performance of a layered security system

Biometric Sample Synthesis. **Figure 4** A photographic image of a live person added to a 3D face model (*Reproduced with permission from the original author*).



Biometric Sample Synthesis. **Figure 5** Methodology for creating synthetic sample databases using parametric models.

that contains security devices that have a known reliability, or can use it to optimize camera locations and lighting conditions.

National security support can be provided in the areas of border control in airports, border crossings, and in ports – by providing the capability to understand potential vulnerabilities through controlled areas. Countries that lack sufficient biometric diversity to test border control systems for under-represented ethnic groups would certainly benefit from the ability to inject synthetic biometrics from these groups to insure the system is not biased in its ability to properly handle those individuals. Systems that under-perform on specific ethnic groups mostly lacked sufficient training data for the engineers building the system.

Synthetic biometric samples are not associated with specific individuals; hence one could then argue that they enhance privacy. Biometric databases generally must be encrypted and secured for protection of an individual's identity, especially when additional metadata about that individual providing the biometric sample is accumulated in the same data-gathering exercise. Synthetic biometric samples do not have this restriction.

A sensitive subject for some is the use of biometrics by governments. The U.S. Department of Defense's anti-terrorism total information awareness system attracted significant congressional and public scrutiny concerning the privacy, policy and potential abuses of a system whose intended purpose was to protect U.S. citizens from individuals known to want to cause the U.S. harm. A concern about the potential ultimately led to the cancellation of the program and is summarized in a December 2003 audit report from the Inspector General of the U.S. Department of Defense [6].

There should be few if any restrictions on sharing the parameters used to create synthetic biometric samples or entire synthetic biometric databases. Further, assuming the modeling science can progress to an advanced state, the engineers and researchers could eventually create standard models, from which they would only need to exchange parameter settings to allow anyone to recreate specific or statistically similar synthetic biometric samples.

Another benefit to using synthetic biometrics is the cost and time savings from the need to acquire real biometric samples for testing systems. Provided the device acquisition and the impacts from factors like environmental changes are model-able, and the effects of presentation variations are well understood, realistic synthetic samples can be quickly generated. The synthetic samples can subsequently be used to augment or perhaps someday reduce the need for system scenario tests, saving money.

As with a number of technologies, synthetic biometrics generators have the potential for misuse. Among these uses are as rapid "hill-climbing" biometric generation devices that can be used to identify people in a biometric system that has not taken appropriate security safeguards to thwart hill-climbing attacks. Another potential misuse would be to characterize an individual's biometric with specific parameters, which could then be used to generate specific synthetic biometrics that could fool biometric systems across a wide variety of possible sensors and environmental conditions through the creation of phony biometrics. Fortunately, biometric system engineers are cognizant of these potential security vulnerabilities and routinely take appropriate precautions to counter potential attacks from phony biometrics [7, 8].

Summary

Biometric technology becoming a ubiquitous addition to many modern security technologies. The synthesis of biometric samples has important benefits that may one day play an important role in the future of biometrics. The likelihood that image biometric sample synthesis of facial or body characteristics may become nothing more than a scientific curiosity is remote. This is due to the movie industry's quest to create lifelike animated avatars.

The biometrics industry lacks validated models. This shortfall remains one of the primary issues facing the use of synthetic biometrics. In addition, the accurate transformation of a specific synthetic biometric between sensors and environments remains as an important next step that has been achieved to a certain degree by at least some of the vendors of these products.

The ultimate potential for synthetic biometrics is providing a cost-effective method to avoid widely publicized biometric deployment failures. The poster child deployment failure was the Boston Logan Airport's attempt to utilize a face recognition system that according to reports failed to match the identities of 38% of a test group of employees. Had the deployment specifics (lighting conditions, algorithms, camera type, angles, etc.) been checked in the lab with a synthesized environment with injected real and synthetic biometric avatars, it is entirely possible that this snafu could have been avoided [9].

Despite some potentially negative uses, there are significant potential benefits from biometric sample synthesis. Increases in sophistication, reliability, and accuracy of synthetic biometrics will improve the potential for decreasing false match and false nonmatch rates in systems through the use of finely tuned biometric samples to allow algorithm improvements to account for numerous noise inducing factors. This improvement would be cost effective and privacy enhancing – provided the synthetics accurately reflect

what a real subject's biometric would (or could) appear like to the system's biometric template extraction and matching algorithms.

Biometric sample synthesis is a technology with promising applications – the potential of which has not been fully realized.

Related Entries

- ▶ Attack trees
- ▶ Biometric security threat
- ▶ Biometric vulnerabilities, Overview
- ▶ Face sample synthesis
- ▶ Fingerprint sample synthesis
- ▶ Iris sample synthesis
- ▶ Markerless 3D Human Motion Capture from Images
- ▶ SfinGe
- ▶ Signature sample synthesis
- ▶ Voice sample synthesis

References

1. Orlans, N.M., Buettner, D.J., Marques, J.: A survey of synthetic biometrics: Capabilities and benefits. Proceedings of the International Conference on Artificial Intelligence (IC-AI'04) 1, pp. 499–505 (2004)
2. Greenberg, D.P., et al.: A framework for realistic image synthesis. Computer Graphics Proceedings of SIGGRAPH 1, pp. 477–494. New York (1997)
3. Buettner, D.J., Orlans, N.M.: A taxonomy for physics based synthetic biometric models. In Proceedings of the Fourth IEEE Workshop on Automatic Identification Technologies (AUTOID'05) 1, pp. 499–505 (2005)
4. Daugman, J. The importance of being random: Statistical principles of iris recognition. Pattern Recogn. **36**, 279–291 (2003)
5. Mansfield, A.J., Wayman, J.L.: 2002. Best Practices in Testing and Reporting Performance of Biometric Devices Version 2.01. *Centre for Mathematics and Scientific Computing, National Physical Laboratory (NPL Report CMSC 14/02)*:10. http://www.npl.co.uk/upload/pdf/biometrics_bestprac_v2_1.pdf
6. Department of Defense Office of the Inspector General (Information Technology Management), Terrorism Information Awareness Program (D-2004-033). December (2003)
7. Soutar, C.: Biometric System Security, p. 4. http://www.bioscrypt.com/assets/documents/whitepapers/biometric_security.pdf
8. Roberts, Chris.: Biometric Attack Vectors and Defences. Computers & Security. **26**(1), 14–256 (2007)
9. Murphy, S., Bray, H.: Face recognition devices failed in test at Logan. *The Boston Globe*. http://www.boston.com/news/ local/articles/2003/09/03/face_recognition_devices_failed_in_test_at_logan. Accessed 3 September 2003

Biometric Security Measure

Biometric security measure (or security countermeasure) is a technological or procedural system designed to protect a biometric system from active attack (Biometric security threat). Examples of security measures include: liveness detection which is designed to detect spoof biometric samples; and cancelable biometrics and biometric encryption which are designed to protect against attacks on Biometric template security. Examples of procedural measures include surveillance and supervision of sensors. Biometric security measures are not designed to defend from zero-effort impostors; as this aspect of the biometric system would be considered the biometric performance.

- ▶ Security and Liveness, Overview

Biometric Security, Standardization

Greg Cannon[1], Philip Statham[2], Asahiko Yamada[3]
[1]Crossmatch Technologies, Palm Beach Gardens, FL
[2]Biometrics Consultant, Specialising in Standards and Security Cheltenham, Gloucestershire, UK
[3]Advanced IT Laboratory, Toshiba Solutions Corporation

Synonym

ACBio instance

Definition

Biometrics holds out the promise of increased confidence in personal authentication processes compared with traditional passwords and tokens (e.g., keys and cards). This is because of the direct link between the biometric characteristic and the individual (strong binding) compared with the indirect link represented by passwords and tokens (weak binding).

Biometric Systems are IT systems that include biometric recognition functionality. The security of

biometric systems shares much with the traditional IT system security, but there are some factors that are biometric specific. These include threats such as spoofing and the personal nature of biometric data that require special handling.

The earliest work on biometric security standards related to biometric security management for the financial services sector. However the recent growth in the deployment of biometric systems, particularly in public domain applications such as passports, visas, and citizen cards, has given a strong impetus to the development of standards that address the comprehensive requirements of biometric systems and application security. Consequently, there is now a concerted effort by the two major standards groups involved ISO (International Organization for Standards)/IEC JTC 1 (Joint Technical Committee 1 (The IT Standards Committee of ISO)) SC37 (Biometric Standards Subcommittee of JTC 1) and SC 27 (IT Security Standards Subcommittee of JTC 1) to cooperate to develop the new guidelines and standards needed to deploy biometric systems securely in the modern world.

Current areas of study include

1. Biometric security evaluation.
2. Biometric transaction security.
3. Protection of biometric data.
4. Guidance for specifying performance requirements to meet security and usability needs in applications using biometrics.

Introduction

The rapid growth of biometric technology for authentication in public domain applications, such as passports, visas, border control and citizen card schemes, is fuelling an intensive program of work to develop international standards for future biometric systems. The availability of standards provides suppliers with a set of specifications and "good practices" as targets for their products, and gives system designers more confidence that their systems will work as intended and be interoperable with other products designed to conform to the same standards. Alongside the technical standards, corresponding security standards are needed to ensure that biometric applications can be designed, built, and deployed with adequate protection for the system and for its users.

Since biometric systems are also IT systems, the threats to security will share some aspects with those of IT systems generally. However, there are specific considerations for biometric systems that lie outside the normal. These include areas such as vulnerabilities, which include the threat of spoofing with an artifact (e.g., gummy fingerprint), mimicry, the capture and replay of biometric data, and privacy concerns because of the personal nature of biometric data. Function creep and identity theft are examples of possible misuse that are particularly relevant to biometric applications. The consequence of these special factors is that, for biometric systems, security considerations need to extend beyond the system to include protection for the biometric data of individuals whose biometric data are processed by or stored on the system.

Although there is already a rich vein of IT security standards available that are applicable to biometric systems, the use of biometrics brings new, biometric-specific, security requirements that existing standards do not deal with. Biometric and IT security standards bodies are currently focused on the development of new biometric security standards that address the deficiencies.

The biometric and IT security standards communities need to collaborate closely because of the vital need for shared expertise and also because of the inevitable dependencies between standards specifying the technology and others aimed at security. For example, providing end-to-end security assurance of biometric transactions across a network will require security information to be generated and processed by the biometric hardware and software at each end of the connection as well as on the linking network. These end-points are governed by the technical biometrics standards BioAPI (Biometric Application Programming Interface) [1] and CBEFF (Common Biometric Exchange Format Framework) [2] developed by SC 37, and these have strong interdependencies with ACBio (Authentication Context for Biometrics) [3], the biometric transaction security standard under development in SC 27. This and other examples are discussed in more detail in later sections of this article.

Close liaison between SC 37 and SC 27 has existed since the formation of SC 37 in December 2002. Each subcommittee has appointed liaison officers who attend meetings of both the subcommittees and take responsibility for identifying projects requiring

cooperation between SC 27 and SC 37 and ensuring that relevant experts can be provided to support them. Recent action taken by SC 37 will further strengthen the cooperation with SC 27 through a coordinated support group operating within SC 37. The motivation is not only for the reasons given earlier but also because much of the biometrics expertise essential for the biometric security standards work is concentrated in SC 37.

The following sections of this article provide a brief discussion of the biometric security issues currently being addressed by the standards community and the associated standards development projects. Readers should however note that, although the information here was correct at the time of writing, many of these standards are still in development and are evolving rapidly; in consequence, some of the information will inevitably become out of date. Readers are therefore urged to visit the web sites of the relevant international standards subcommittees for the current status of biometric security standards. The URLs are listed in the reference section at the end of the article [4, 5].

Biometric Data Management Standards

Biometric Data Management is concerned with the broader issues of management and procedural measures for protecting biometric data. These include awareness training and accounting and auditing procedures as well as a reference to technical measures such as those described in this article.

Historically, this work originated from the ANSI X9 subcommittee in the US X9.84 Standard – Biometric Information Management – Security (2003) [6]. X9.84 progressed into the international standards domain to become the starting point for the development of ISO 19092-2008 – Financial services – Biometrics – Security Framework [7]. ISO 19092 is a biometric-specific extension of ISO 17799, the Code of Practice for Information Security Management, which is now subsumed into the ISO 27000 family of ISMS standards [8].

Biometric Data Security Standards

Biometric data stored and processed in biometric systems are security sensitive. Their loss or disclosure could potentially lead to the undermining of the authentication integrity of the system and misuses such as function creep, identity theft, and breaches of personal privacy. The disclosure of biometric reference data (e.g., fingerprint templates) might provide identifying information for an attacker to transfer to an artifact for use in a spoofing attack, or to generate an electrical signal that could be directly injected in an electronic attack. If exported for use elsewhere without the authority of the individual, this would constitute function creep and possibly a breach of privacy. In many countries, such practices are regulated by data protection legislation or codes of conduct.

To guard against these threats, various procedural and technical measures can be employed. Current technical standards work focuses on the protection of stored biometric data, including biometric samples and biometric references, using cryptographic techniques such as digital signatures and encryption.

The core standard for biometric data storage and exchange is ISO/IEC 19785 CBEFF (Common Biometric Exchange Format Framework). CBEFF is a multipart standard where Part 4–Security block format specifications–provides for the protection of biometric data integrity and confidentiality.

The CBEFF standard defines a basic block of biometric data called a BIR (Biometric Information Record). The BIR is further subdivided into a Standard Block Header (SBH), a Biometric Data Block (BDB) containing the biometric data themselves (which may be encrypted), and a Security Block (SB). The SBH header includes indicators of the security mechanisms that are used to protect the data. The SB security block contains relevant security information such as cryptographic checksums, digital certificates, and data encryption algorithm specifications etc. that are used to guarantee the integrity and confidentiality of the data. The details of these options and the structure of SB are being standardized in 19785-4 CBEFF Part 4, using The Internet Society's RFC 3852 CMS (Cryptographic Message Syntax) [9]. The specifications within the CBEFF security block are planned to encompass the security requirements associated with the ACBio (Authentication Context for Biometrics) standard [3], which is being developed in SC 27 to provide end-to-end assurance for biometric transactions. Essentially, the CBEFF security block will contain a set of ACBio instances which contain data that can be used to validate the end-to-end integrity of the biometric

transaction. Further information on ACBio appears in the next section of this article.

SC 37 biometric standards are being modified in order to support ACBio. The effect on CBEFF has been described, but the BioAPI (ISO/IEC 19784-1 Information technology – Biometric application programming interface – Part 1: BioAPI specification) is also in the process of being updated to accept BIRs, including Security Blocks. An Amendment 3 to the BioAPI standard is under development to deal with the extended requirement for the provision of security of data.

One approach to protecting biometric data is to replace the central database of biometric references by storage of each enrollee's reference on a personally held smartcard. This is often advocated by groups concerned about the privacy implications of centralized biometric databases. Secure smartcards could also provide the necessary biometric processing, the main system capturing the biometric sample, passing the sample to the smartcard for matching against the reference stored on the card, and authenticating the result delivered by the smart-card. This is what is known as "On-card matching". A claimant could carry the smartcard with him/her; present the card to the system together with a biometric sample; and assure the system that he/she is genuine by allowing the secure processor of the smartcard to perform the comparison between the live sample and the stored reference. In this way, the biometric data and the comparison algorithm are immune from attacks on the central system.

The SC 37 19794-2 Fingerprint Minutia Standard includes a section specifying a compact format fingerprint minutiae standard suitable for the limited storage capability of smartcards. We envision that more standards may be necessary, especially standards that allow for more interoperability between the smartcard and the IT system.

Biometric Transaction Security Standard – ACBio

Transaction security standards are well established in the IT world, principally driven by the banking and financial sectors where transactions need to be secure not only over private networks but also between banks and customers using the Internet. These standards typically involve secure protocols using digital certificates and data encryption to guarantee the integrity and confidentiality of remote transactions. If transactions are to include biometric authentication, the security envelope needs to extend to provide assurance for the biometric elements of the transaction. Such assurance might include the authentication of the biometric hardware (e.g., fingerprint reader), certification of biometric performance capability, the quality of the current biometric authentication instance, and the integrity of the biometric data transfer process.

This is the scope of the SC 27 standard 24761 Authentication Context for Biometrics (ACBio) [3]. ACBio specifies the structure of data that can provide the necessary assurance for a remote biometric verification transaction.

ACBio models a biometric transaction as a set of processes executed by Biometric Processing Units (BPUs). A BPU places relevant security data into a block called an ACBio instance. BPUs generate and transmit ACBio instances together with the associated biometric transaction data. ACBio instances secure the integrity of the data, using security techniques such as digital signatures and cryptographic checksums. ACBio instances can also contain data that provide the means of assuring other aspects of the transaction such as validation of the biometric hardware used and the certification of the performance capability of the biometric verification process.

Transactions passing between BPUs will typically accumulate a collection of ACBio instances associated with the various processing stages. Each ACBio instance will contain security markers (cryptographic checksums, digital signatures etc.) that can provide assurance for the corresponding process stages. Further details are beyond the scope of this article, but the security techniques used can provide protection against the substitution of "bogus" components and data replay attacks as well as general threats to the integrity of the transaction data.

ACBio instances depend on other biometric and security standards for their operation and effect. Interdependencies with the CBEFF and BioAPI standards have already been described in the *Biometric Data Security Standards* section. Other standards are also referenced by ACBio. An ACBio instance uses data types defined in the RFC 3852 CMS (Cryptographic Message Syntax) standard [2]. ACBio instances also use X.509 digital certificates [10]. For the certification

of biometric performance capability, ACBio calls on the SC 37 19795 series of biometric performance test standards [11]. To provide test results in a suitable format for use by ACBio, work has begun in SC 37 on the 29120 standard: Machine-readable test data for biometric testing and reporting [12]. Work is also expected in SC 27 to produce a standard for the electronic format of cryptographic modules that will be used by ACBio. Finally, ACBio refers to the SC 27 19792 Biometric Evaluation Methodology standard [13] to provide security assurance for the biometric hardware and software used in an application.

ACBio will therefore use existing cryptographic and digital certificate techniques to assure transaction data integrity end-to-end. The integrity of the biometric hardware and the performance and security of the biometric technology will be provided by external evaluation schemes, and the results will be embedded in machine-readable data formats that can be authenticated by the validation of the biometric verification process as required.

The multiple dependencies between SC 27 and SC 37 standards for the successful operation of ACBio call for close ongoing cooperation between the two subcommittees to ensure consistency and interoperability of the standards. Other collaborations are also required. In the area of smart cards, there is collaboration between SC 17 and SC 27 to include in ACBio an informative annex of command sequences for the realization of ACBio on STOC (STore On Card) cards and OCM (On Card Matching) cards. A STOC card is a smart card that stores the biometric reference data on the card, but does not perform the biometric verification, and an OCM card is a smart card that both stores biometric reference data and performs the biometric comparison between the reference and the input biometric sample data.

Biometric System Security Evaluation Standards

Historical Background

Biometrics is about identification and verification. However, in many systems, failures of identification or verification will have security implications. Often the reason that biometric technology is used is because of the perceived increase in assurance of correct identification or verification that biometrics will provide. However, to reliably assess this level of assurance, a properly constituted security evaluation procedure is needed.

Security evaluation of IT systems is now well established. Various evaluation schemes exist for specific market sectors such as antivirus products and smartcards. The internationally recognized standard for IT security evaluation is ISO 15408 – Common Criteria [14]. This is a government-developed scheme aimed primarily at evaluation for government use, but it is also recognized and used commercially as a "gold standard" for security evaluation. Evaluations are performed by government-licensed evaluation laboratories in member countries and the results are recognized across the participant countries (and wider) through a mutual recognition agreement.

Although the Common Criteria evaluation methodology is generic and therefore suitable for biometric system evaluations, there are a number of special factors that need to be considered when undertaking biometric system security evaluations. These include statistical performance testing and biometric-specific vulnerabilities. This was first recognized during a pioneering Common Criteria evaluation of a biometric fingerprint verification system in Canada in 2000 [15], which led the evaluation team to investigate and develop the methodology to deal with the special factors. Subsequently, this work was further developed by an informally constituted group of biometric and Common Criteria experts to produce a biometric evaluation addendum for the Common Criteria Methodology known as the Biometric Evaluation Methodology or BEM [16]. The BEM describes the special requirements of a biometric system security evaluation and gives guidance to evaluators on how to address these requirements in a Common Criteria evaluation. At the time of writing, the BEM had not attained official status as a formal part of CC methodology. Nonetheless, it is frequently referenced as a source of information on CC and other security evaluations of biometric products and systems.

ISO/IEC 19792: Information Technology – Security Techniques – Security Evaluation of Biometrics [13]

This international standard is currently under development in SC 27. Project 19792 is not targeted at a

specific evaluation scheme such as Common Criteria; rather, its aim is to provide guidance to developers and evaluators on security concerns for biometric systems and to specify a generic methodology for their evaluation. It is similar to the BEM, but is not limited to Common Criteria evaluations and contains more detailed information on potential threats, countermeasures, and evaluation requirements. Like the BEM, it assumes that evaluators are familiar with the broader IT security evaluation issues and does not address these.

19792 covers biometric-specific security issues of the system as a whole as well as threats and potential vulnerabilities of the component parts. It describes technical and nontechnical threats and how these may be reduced or eliminated by appropriate countermeasures. It provides guidance to evaluators on testing and the assessment of potential vulnerabilities and countermeasures, and it defines the responsibilities of vendors and evaluators in the evaluation process.

Biometric-specific aspects of system security and evaluation methodology covered by 19792 include.

Statistical Performance Testing

Biometric comparison decisions (match and nonmatch) are not certainties, but are prone to false match and false non-match errors. Comparison results are therefore often expressed in terms of the probabilities of correct and incorrect decisions, the actual numbers being expressed in terms of statistical performance figures. An example of what this means in practical terms is that for an access control application with a false match rate of 1%, if 100 randomly chosen impostors were to present their own biometric characteristic to the system while claiming to be legitimate enrollees, one of them might succeed in gaining admittance through chance error. The quantification of errors through robust performance testing therefore forms one part of a biometric system security evaluation. The international standard for biometric testing and reporting is provided by the multipart ISO/IEC standard 19795 [11].

The significance of biometric error rates to security depends on the purpose of the identification or verification in the application domain. For access control, the false match rate may be the most important security relevant factor, but for applications such as passport or ID card registration, an important requirement will be the successful detection of attempts to register multiple times under different claimed identities. Here, the system needs to search its biometric database to determine if there is an apparent match with any existing enrollee. If a false non-match occurs during the search, a multiple enrolment attempt may succeed and therefore, for this function, the false non-match rate statistics will be the most important security consideration.

Biometric System Threats and Countermeasures

The use of biometrics brings potential security threats and vulnerabilities that are distinct from those of other IT technologies, including spoofing, mimicry, and disguise. Further details of these threats and examples of countermeasures can be found in the definitional entries for ▶ Biometric System Threats and ▶ Countermeasures.

Human Security and Privacy Concerns

Since biometric systems collect and store the personal data of its enrollees, security measures are necessary to protect the data and the privacy of the enrollees. This is another important difference between systems using biometrics for authentication and those that depend on inanimate entities such as passwords and tokens.

People have a right to privacy regarding the use and sharing of their personal data, that is, data about their lifestyle, preferences, habits etc. that can be linked to them as individuals. Such data should be collected, processed, and stored only with the informed consent of the individual and only for the declared and authorized purpose. Unauthorized disclosure and misuse can lead to undesirable consequences such as identity theft and function creep. Biometric data are regarded as particularly sensitive, because their strong binding to specific persons may make it difficult for individuals to repudiate transactions authorized by biometric authentication.

Technical security measures such as data encryption and the use of cryptographic signatures to bind data to an application can help to secure biometric data, but usually, complete protection also requires administrative controls and sanctions implemented within an overall system security policy.

Future Directions for Biometrics Security Standards

The first generation of biometric standards may be characterized as a collection of largely self-contained or stand-alone parts that provide the essential building

blocks for future biometric systems. These building blocks are now largely in place, but the course of their development has uncovered new areas of work that need to be addressed by a second generation of biometric standards.

Building on the experience of developing the earlier standards, the second generation will target the broader requirements for system and application level standards. The new standards will tackle areas that were omitted from the first generation standards and serve to bind together the earlier work to furnish a comprehensive standards package that will meet the wider systems and applications level standards requirements. Biometric system designers and implementers need these standards to support the rapid growth in large public domain biometric systems that we are now seeing, including passports, visas, border control applications and financial transaction systems. Many of these systems are international in reach and raise important privacy and other human concerns as well as major technical challenges.

In the security area, work is needed on standards that deal with such issues as

1. The use of multimodal biometrics to increase the security that biometric authentication offers;
2. Comparing and quantifying the security capabilities of biometrics and password- and token-based authentication technologies individually and in combination;
3. Assessing the requirement for biometric performance in the context of a system where biometrics provides only one element of security as part of an overall system security policy;
4. The potential role of biometric authentication in identity management systems;
5. Locking biometric data to specific applications to prevent misuse and potential identity theft;
6. Referencing, interpreting, and using other relevant security standards, for example, US Government Federal Information Processing Standards FIPS 140 for data encryption; X.509 digital certificates, in the domain of biometric security standards.

Some groundwork has already begun. In the United States, the InterNational Committee for Information Technology Standards (INCITS) M1 Standards Committee has picked up on earlier work by the US National Institute of Standards and Technology (NIST) on Electronic Authentication and E-Authentication for US Federal Agencies [17, 18] and produced a study report on the use of biometrics in e-authentication [19].

A special group has been formed by SC 37 to study and develop a proposal for future work on providing guidance for specifying performance requirements to meet security and usability needs in applications using biometrics. Both this initial study and any subsequent work will require close cooperation and involvement of experts from other standards subcommittees, in particular, SC 27.

Related Entries

- ▶ Biometric Technical Interfaces
- ▶ International Standardization
- ▶ International Standardization Finger Data Interchange Format
- ▶ Performance Testing Methodology Standardization

References

1. ISO/IEC JTC 1 SC 37 19784 Biometric Application Programming Interface (BioAPI). Multi-part standard, some parts under development at the time of writing
2. ISO/IEC JTC 1 SC 37 19785 Common Biometric Exchange Format Framework (CBEFF). Multi-part standard, some parts under development at the time of writing
3. ISO/IEC JTC 1 SC 27 24761 Authentication Context for Biometrics (ACBio). Standard under development at the time of writing
4. SC 27 http://www.iso.org/iso/standards_development/technical_committees/list_of_iso_technical_committees/iso_technical_committee.htm?commid=45306. Accessed 30 October, 2007
5. SC 37 http://www.iso.org/iso/standards_development/technical_committees/list_of_iso_technical_committees/iso_technical_committee.htm?commid=313770. Accessed 30 October, 2007
6. ANSI X9.84 Biometric Information Management and Security for the Financial Services Industry see: http://www.techstreet.com/cgi-bin/detail?product_id=1327237 for further details. Accessed 30 October, 2007
7. ISO 19092-2008 – Financial services – Biometrics – Security Framework. ISO 19092-1 see: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=50145 for further details. Accessed 30 October, 2007
8. ISO 27000 family of Information Security Management Systems (ISMS) standards see: http://www.itgovernance.co.uk/infosec.aspx for further details. Accessed 30 October, 2007
9. RFC Cryptographic Message Syntax 3852. The Internet Society – see ftp://ftp.rfc-editor.org/in-notes/rfc3852.txt. Accessed 30 October, 2007

10. ITU-T X.509 Information technology – Open Systems Interconnection – The Directory: Public-key and attribute certificate frameworks
11. ISO/IEC JTC 1 SC 37 19795 Biometric Testing and Reporting. Multi- part standard, some parts under development at the time of writing
12. ISO/IEC JTC 1 SC 37 29120 Information Technology: Machine Readable Test Data for Biometric Testing and Reporting. Multi-part standard under development at the time of writing
13. ISO/IEC JTC 1 SC 27 19792: Information technology – Security techniques – Security evaluation of biometrics. Standard under development at the time of writing
14. ISO/IEC 15408, Common Criteria for Information Technology Security Evaluation, http://www.commoncriteriaportal.org/
15. Bioscrypt™ Enterprise for NT Logon, version 2.1.3: Common Criteria Evaluation http://www.cse-cst.gc.ca/services/cc/bioscrypt-eng.html. Accessed 30 October, 2007
16. Common Criteria. Common Evaluation Methodology for Information Technology Security Evaluation – Biometric Evaluation Methodology Supplement (BEM) http://www.cesg.gov.uk/policy_technologies/biometrics/media/bem_10.pdf
17. NIST SP800-63, Electronic Authentication Guideline, http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf. Accessed 30 October, 2007
18. OMB M-04-04, E-Authentication Guidance for Federal Agencies, http://www.whitehouse.gov/omb/memoranda/fy04/m04-04.pdf. Accessed 30 October, 2007
19. INCITS M1/06-0642 Study Report on Biometrics in E-Authentication, http://m1.incits.org

Biometric Security Threat

Biometric Security Threat is an approach of active attack against vulnerability in a biometric system (see Biometric system: vulnerabilities). Threats may be broadly classified as: Presentation attacks (spoofing), in which the appearance of the biometric sample is physically changed or replaced; Biometric processing attacks, in which an understanding of the biometric algorithm is used to cause incorrect processing and decisions; Software and networking vulnerabilities, based on attacks against the computer and networks on which the biometric systems run; and Social and presentation attacks, in which the authorities using the systems are fooled. To defend against a biometric security threat, a biometric security measure may be used.

► Security and Liveness, Overview

Biometric Sensing

► Biometric Sample Acquisition

Biometric Sensor and Device, Overview

Geppy Parziale
iNVASIVE CODE, Barcelona, Spain

Synonyms

Biometric sensors; Biometric devices

Definition

A biometric sensor is a transducer that converts a biometric treat (fingerprint, voice, face, etc.) of a person into an electrical signal. Generally, the sensor *reads* or *measures* pressure, temperature, light, speed, electrical capacity or other kinds of energies. Different technologies can be applied to achieve this conversion using common digital cameras or more sophisticated combinations or networks of sensors.

It is important to highlight that the output signal of a sensor or device is only a representation of the real-world biometrics. Hence, if B is a biometrics of a real-world and s is the transfer function of a sensor or a device, the output signal is $B' = s(B)$ and $B' \neq B$.

A biometric device is a system which a biometric sensor is embedded in. Communication, processing and memory modules are usually added to provide additional functionalities that the biometric sensor cannot if standalone.

Interchanging the terms *sensor* and *device* is very a common practice, even if they are two different concepts. A sensor is responsible only for the conversion of a biometrics into an electrical signal. Instead, when a processor and a memory module are also involved, the term device is more appropriate.

Introduction

Biometrics identification and verification are slowly penetrating the security market. The convenience of avoiding to recall passwords and/or loose tokens (id cards, smart-cards, etc.) is one of the strongest advantage of biometrics compared to the legacy security tools. Moreover, biometrics concretely links a person to her/his identity, compared to the traditional approaches that associate the person identity to a token or a password that can be forged, lost, forgotten or used by other people.

The most sensitive step in the biometric authentication chain is the ▶ biometric capture. The accuracy and the repeatability of this process influence the remaining steps of the chain. Since the output signal of a device is only a representation of the real-world biometrics, the choice of the representation type is a very important issue, because it should try to meet the four biometric axioms: uniqueness, repeatability, permanence and collectibility [1]. However, this is a very complex task influencing the choice of a technology used and the design of the sensor/device for a defined application.

Biometrics sensors must be designed taking into account many factors. User convenience, portability, electrical and optical characteristics and price are only some of them. They are very important factors when choosing among different sensors for a defined application. However, they cannot be always met and the right balance of these factors has to be found according to the final application in which a biometric sensor will be involved.

Below, the main features of a biometric sensor and device are reported. This is not intended to be an exhaustive list of features and only the most important characteristics are highlighted.

User Acceptance

User acceptance is an important factor that has to be taken into account during the choice of a technology and the design of a biometric device. Easy-to-use devices are preferable to user-unfriendly ones. For example, devices pointing lasers to the eyes or providing small electrical current to the body of a person are for sure difficult to be accepted by the final user. Sensors touching a person body are less preferred than remote sensors or device re-used to touch many individuals are not well accepted for hygienic reasons. Some biometric devices could be difficult to be accepted because of cultural or religious motivations.

In general, biometrics sensors and devices can be classified in two main families: *intrusive* and *non-intrusive*. The closer the device to the person, the more intrusive the device. For example, a surveillance camera able to identify people face remotely is less intrusive than imaging sensors touching the user eye to scan the retina.

Some biometric devices need the user to cooperate during the capture and offer her or his own biometrics. Other devices do not need any user cooperation. Moreover, when an operator is needed during the normal use, the device can be classified as a *supervised device*, while when the user can operate the sensor with no extra support, it can be classified as an *unsupervised device*. Usually, unsupervised devices are preferred to supervised ones, because they do not need extra human resources to operate.

Portability

Form factor and weight are sometimes very important characteristics that must be taken into account during the design of a biometric sensor, because they can influence its portability. Embedding a face or fingerprint or iris sensor (or all together) in a mobile handheld computer or laptop or cellphone is becoming a very attractive solution for different kinds of applications. When the portability is important, the sensor is usually embedded in a more complex device containing all the functionalities (signal processing, communication, matching, etc.) that the biometrics sensor cannot provide alone. The possibility to process locally the captured biometrics requires the existence of processor and memory modules. Instead, when the processing is performed remotely a communication interface must be considered as part of this more complex device. In both cases, the power consumption becomes an issue, because the need of supplying the energy through portable batteries can limit the choice of the technology.

Ruggedness and Lifetime

When a biometric device has to be installed or carried in difficult environments (very low or high temperature,

high humidity, vibrations, dust, noisy locations) or when mechanical moving parts are involved (line-scan cameras, auto-focusing cameras, auto-position sensors, etc.) important features are the ruggedness and the lifetime. These influence the maintenance costs of the device. Thus, during the design, these features have to be taken into account and special housing or materials must be used for the sensor manufacturing.

Calibration

The standard functionality of a sensor is usually influenced by the external or internal factors and thus, it can change during time, due to temperature, pressure or humidity variations or due to some mechanical movements. To reduce this problem, sensors need to pass a periodical procedure to restore the initial operational conditions. This process is called sensor or device ▶ calibration.

Calibration refers to different processes used according to the type of sensor and the technologies involved for the capture. Electrical calibration is the process used to restore the initial electrical conditions that could change over time due, for example, to temperature variations. Mechanical calibration is performed instead when a device has moving parts. In this case, mechanical frictions starts to appear during the normal sensor life altering the measure the sensor was designed for. Optical calibration is instead the process used to re-focus lenses or re-establish the initial illumination conditions.

The calibration is sometimes a process that is also needed when the sensor is used for the first time (out-of-the-box). Due to inaccuracies of the manufacturing, the sensor functionality can be slightly different than the defined one. Positioning, orientation or placement of sensor parts can be sometimes very difficult and the production process are usually not free of imperfections. Thus, the first time the device is used and then periodically, a calibration procedure is needed. This can be a manual, semi-automatic or fully-automatic procedure. Fully-automatic calibration is usually possible when the biometric device does not contain mechanical and optical parts. In this case, the sensor calibration is usually obtained using special electrical circuits controlling the status of the device and re-establishing the correct initial electrical conditions. Optical calibration often requires the use of special ▶ optical targets. These are mechanical models used to measure pre-defined known values against which the output of the sensor is compared. Mechanical calibration is usually done manually by an experienced operator, reviewing all the mechanical functionalities.

Operating Conditions

The set of conditions (e.g., voltage, temperature, humidity, pressure, etc.) over which specified parameters maintain their stated performance rating are called *operating conditions*. When these are not respected, the biometric sensor could not work as defined by the manufacturer. The operating conditions must be chosen according to the final sensor applications. Sensors used for military applications have usually very large operating conditions and the devices is supposed to work under huge stress (high or low temperatures, vibrations, dust, high humidity, etc.). As other electrical or mechanical components, biometric sensors must meet some standard requirements and pass a certification process. For example, ▶ ISO certifications define the electrical and mechanical characteristics that an electronic device should meet to be sold.

Sensor Interface

The possibility to interface a sensor or device with other sensors or devices and with a processing unit is an important feature that must be considered when choosing a sensor for a defined biometric application. USB and Firewire can be the best choice, when the biometric sensor needs to be connected to a standard PC. When the data throughput is an issue, optical fibers or gigabit ethernet are possible solutions. Moreover, if the quantity of data the sensor has to transfer to a processing unit is large, the interface must be able to transfer this data as fast as possible to avoid long latency. Wired or wireless communication interfaces can be chosen according the final application.

Power Supply

Low-current absorption is usually a very required feature for a biometric device, because this facilitate to embed it in other devices. Usually the basic sensors (e.g. cameras and microphones) do not need to drain

too much current, but when illuminators (▶ Light Emission Diode or optical fibers) or mechanical movements (line scanning cameras) or heating generators (palmprint devices reducing the halo effect) are involved, then extra power is needed. Modern communication interfaces as USB 2.0, Firewire and Ethernet can supply power to the sensor with no need of extra wires. This is a very interesting alternative especially when the biometric application requires a portable device connected to a laptop.

Failure Rate

Failure Rate is the frequency with which an engineered system or component fails. It can be expressed in failures per hour. Mean Time Between Failures (MTBF) is the mean (average) time between failures of a system, and is often attributed to the *useful life* of the device i.e., not including *infant mortality* or *end-of-life*, if the device is not repairable. Calculations of MTBF assume that a system is renewed, i.e. fixed, after each failure, and then returned to service immediately after failure. The average time between failing and being returned to service is termed mean-down-time (MDT) or mean-time-to-repair (MTTR).

Cost

The cost of a biometric device or sensor is a very important factor influencing the final target application in which the device or the sensor will be involved. The final costs depend on many factors. The availability of the basic technology used for the biometric capture is one of them. If special and sophisticated technologies are used instead of the common ones, the costs of the device increase. Moreover, the production materials, the manufactured number of samples and the maintenance are also factors influencing the final costs.

Sensor Resolution

Sensor resolution refers to the ability of a device to acquire, scan or distinguish details of the acquired biometric treat. Depending on the sensor type, it can be distinguished among spatial, frequency, time and radial resolution. For example, a face device can be an area-sensor and its spatial resolution measures the quantity of details of the face skin it can acquire.

Spatial resolution represents the number of pixels in a unitary length and is usually expressed in *pixel-per-inch* or shortly, *ppi*. *Frequency resolution* represents the ability of a device to distinguish frequency variations. *Time resolution* measures the ability of a sensor to distinguish time variations. For example, microphones used as speech devices should have a certain capacity to recognize fast speakers. *Radial resolution* represents the ability of a sensor to distinguish variation in the distance.

The increase of the resolution increases the accuracy of the sensor and usually its final cost. In many applications, a trade-off between resolution and final cost must be found.

Optical and Imaging Characteristics

When a biometric sensor generates as output signal an image and an optical system is involved in the capture process, the choice of a sensor is based on optical characteristics.

Image Depth or *Dynamic Range* determines how finely a sensor can represent or distinguish differences of intensity. It is usually expressed as a number of gray levels or bits. For example, 8 bits or 256 gray levels is a typical dynamic range of fingerprint image or 24 bits or 256 Red, Green and Blue (RGB) levels which is typical of face image.

The *Modulation Transfer Function* (MTF) or *Spatial Frequency Response* represent the relationship between the input and the output signal of a sensor. Spatial frequency is typically measured in cycles or line pairs per millimeter (*lp/mm*). The more extended the response, the finer the detail and the sharper the image. MTF is the contrast at a given spatial frequency f relative to contrast at low frequencies and it ca be computed with the following (1):

$$MTF = 100\% \frac{C(f)}{C(0)}, \tag{1}$$

where $C(f) = (V_{max} - V_{min})/(V_{max} + V_{min})$ is the contrast at frequency f and $C(0) = (V_W - V_B)/(V_W + V_B)$

is the low frequency contrast. V_B, V_W, V_{min} and V_{max} represent the luminance for black areas, the luminance for white areas, the minimum luminance for a pattern near spatial frequency f and the maximum luminance for a pattern near spatial frequency f, respectively.

Geometric Image Accuracy represents the absolute value of the difference $D = X - Y$, between the distance X measured between any two points on the input image and the distance Y measured between those same two points on the output image. This is a very important parameter especially for devices having a very large capture area. This feature is measured using special optical targets.

The capacity of a sensor to capture the whole biometrics in a single image is expressed by the *Field-of-View* (FoV). For a digital camera, this represents the angular extent of the observable object that is seen at any given moment. For some biometrics devices, it is fundamental to capture the biometrics in a single capture. For example, hand-geometry devices needs to capture the full hand in a single shot. Sweep fingerprint sensors allow only the capture of a fingerprint in different instant of times, since their FoV is very limited.

Precise focus is possible at only one distance; at that distance, a point object will produce a point image. *Depth-of-Field* (DoF) represents the range of distance in which the object remains focused. This is a very important feature for remote cameras, since it represents the location in which the biometrics must be placed to be always focused.

The *Intensity Linearity* represents the capacity of a device to reproduce the intensity level values correctly. To prove this feature, a target with gradually varying grayscale levels is usually used for this scope. The grayscale levels on the output image are compared with the grayscale levels on the input target to measure the accuracy of the representation. Large variations in the representation lead that the sensor is calibrated.

The *Signal-to-Noise Ratio* is a measure of the level of noise introduced by the sensor during the biometric capture. This is usually measured using a special optical target representing an intensity level as a reference.

The *Framerate* is the number of frames per time unit that a sensor can generate. It is usually measured in *frames/s*. These parameter is very important when the object movements are involved (sweep devices, touchless devices, gait device, face device) during the biometric capture.

In optics, the *F-number* (sometimes called focal ratio, f-ratio, or relative aperture) of an optical system expresses the diameter of the entrance pupil in terms of the effective focal length of the lens; in simpler terms, the f-number is the focal length divided by the aperture diameter. It is a dimensionless number that is a quantitative measure of lens speed, an important concept in photography.

The *Shutter-speed* is the time that a detector needs to capture a single image. In photography, shutter speed is the length of time while the shutter is open; the total exposure is proportional to this exposure time or duration of light reaching the film or image sensor.

Summary

Biometric sensors and devices are slowly penetrating the security market, because of the advantages of biometrics with respect to traditional security means as passwords and tokens. The choice of a sensor for a defined application is usually dependent on some electrical, ergonomic, optical, mechanical and other characteristics. An overview of this important features has been here reported.

Related Entries

- ▶ Authentication
- ▶ Biometric Sample Acquisition Enrollment

References

1. Clark, J., Yulle, A.: Data Fusion for Sensory Information Processing Systems. Kluwer Academic, Boston, MA, USA, (2009)

Biometric Sensors

▶ Biometric Sensor and Device, Overview

Biometric Services

Biometric functions offered and performed by a service provider on behalf of a requester, usually remotely. Biometric services may include biometric data management, 1:1 verification, or 1:N identification services.

► Biometric Interfaces

Biometric Specific Threats

Synonyms

Attacks; Threats; Vulnerabilities

Definition

Spoofing is the use of an artifact containing a copy of the biometric characteristics of a legitimate enrolee to fool a biometric system. Examples include: gummy fingers, photograph of a face or iris pattern, artificial hand, etc., depending on the modality of the biometric characteristic.

Mimicry is imitating someone else's behavior to fool a biometric system that uses human behavior rather than biology as a distinguishing characteristic. Examples include signature and voice recognition.

Disguise is concealing biometric characteristics to avoid recognition. It can apply to biological and behavioral characteristics and may or may not involve the use of artifacts.

Weak algorithms are biometric algorithms designed to work effectively with the normal range of human characteristics that may behave unpredictably when presented with highly abnormal input signals. This could produce much higher error rates than usual for these abnormal cases. Such signals could be introduced through the use of artefacts or electronically injected via signal replay, e.g., fingerprint with an abnormally high (or low) number of minutiae points.

Capture/replay attack is the capture and subsequent replay of signals flowing in a biometric system, either electrically injected or via transfer to an artifact.

If the biometric system returns a score to the user indicating how close a submitted sample is to the matching decision threshold, it may be possible for an attacker to conduct a methodical attack by making small alterations to successively submitted samples, looking to gradually nudge the score until it passes the matching decision threshold. This is a hill climbing attack.

Database attack is the unauthorized access to biometric data held in the system database, may allow an attacker to inject data or transfer it to an artifact to fool the system.

Biometric systems have environmental vulnerability. Abnormal conditions, like lighting, could cause a biometric system to behave unpredictably, possibly leading to high error rates. Knowledgeable attackers could exploit such a weakness by creating adverse environmental conditions.

► Biometric Security, Standardization

Biometric Spoof Prevention

Biometric spoofing is a method of attacking biometric systems where an artificial object is presented to the biometric sample acquisition system that imitates the biological properties the system is designed to measure, so that the system will not be able to distinguish the artifact from the real biological target.

Biometric spoof prevention involves providing the system with measurement and analysis mechanisms that help differentiate the real biological target from various classes of fake targets. There are several approaches to implementing biometric spoof detection and they are:

- Highly detailed analysis of the primary biometric data can detect and reject low resolution spoofs
- Measurement of secondary properties of the biological target can make spoof fabrication more difficult
- Measurement of variation in the biometric property over short time durations can help reject rigid and stationary spoofs
- Simultaneous measurement of a second biometric property of the same biological target can

significantly increase the complexity and difficulty involved in fabricating fake target artifacts.

► Anti-Spoofing
► Biometric Sample Acquisition

Biometric Strength of Function

The strength of security of the biometric system, being measured through the FAR achieved in an operational environment.

► User Interface, System Design

Biometric Subsystem Transaction Time

► Operational Times

Biometric System

The integrated biometric hardware and software used to conduct biometric identification or authentication. Biometrics is the measurement of physical characteristics, such as fingerprints, DNA, retinal patterns, or speech patterns, for verifying the identity of individuals.

► Biometrics
► Multispectral and Hyperspectral Biometrics

Biometric System Components

Elements of a biometric system, including capture, feature extraction, template generation, matching, and decision.

► User Interface, System Design

Biometric System Design, Overview

Anil K. Jain[1], Karthik Nandakumar[2]
[1]Michigan State University, East Lansing, MI, USA
[2]Institute for Infocomm Research, A*STAR, Fusionopolis, Singapore

Definition

Biometric system design is the process of defining the architecture, selecting the appropriate hardware and software components and designing an effective administration policy such that the biometric system satisfies the specified requirements. The requirements for a biometric system are typically specified in terms of six major design parameters, namely, accuracy, throughput, cost, security, privacy and usability.

Introduction

In general, biometric systems consist of seven basic modules that operate sequentially [1] as shown in Figure 1. These building blocks or modules include (i) a user interface incorporating the biometric reader or sensor, (ii) a quality check module to determine whether the acquired biometric sample is of sufficient quality for further processing, (iii) an enhancement module to improve the biometric signal quality, (iv) a feature extractor to glean only the useful information from a biometric sample that is pertinent for the person recognition task, (v) a database to store the extracted features along with the biographic information of the user, (vi) a matcher to compare two feature sets during recognition and to determine their degree of similarity and (vii) a decision module that determines the user identity based on the similarity (match scores) output by the matcher.

Though all biometric systems are composed of the same basic modules, there are three main steps involved in the design of a biometric system. Firstly, the designer needs to choose the appropriate architecture for a biometric system. Secondly, the hardware and software components required for the implementation of the architecture must be selected. Finally, appropriate policies must be defined for the effective administration of the biometric system. Before the essay dwells

deeper into these three issues, it is important to remember that the goal of any design process is to develop a system that satisfies the *requirements* of the *application*. Hence, most of the design decisions in a biometric system are fundamentally driven by the nature or functionality of the application and the specified requirements.

The functionalities provided by a biometric system can be broadly categorized as *verification* and *identification* [2]. In verification, the user claims an identity and the system verifies whether the claim is genuine by comparing the input biometric sample to the template corresponding to the claimed identity. In identification, the user's biometric input is compared with the templates of all the persons enrolled in the database and the system returns either the identity (in some scenarios, multiple identities whose templates have high similarity to the user's input may be returned by the system.). Of the person whose template has the highest degree of similarity with the user's input or a decision indicating that the user presenting the input is not an enrolled user.

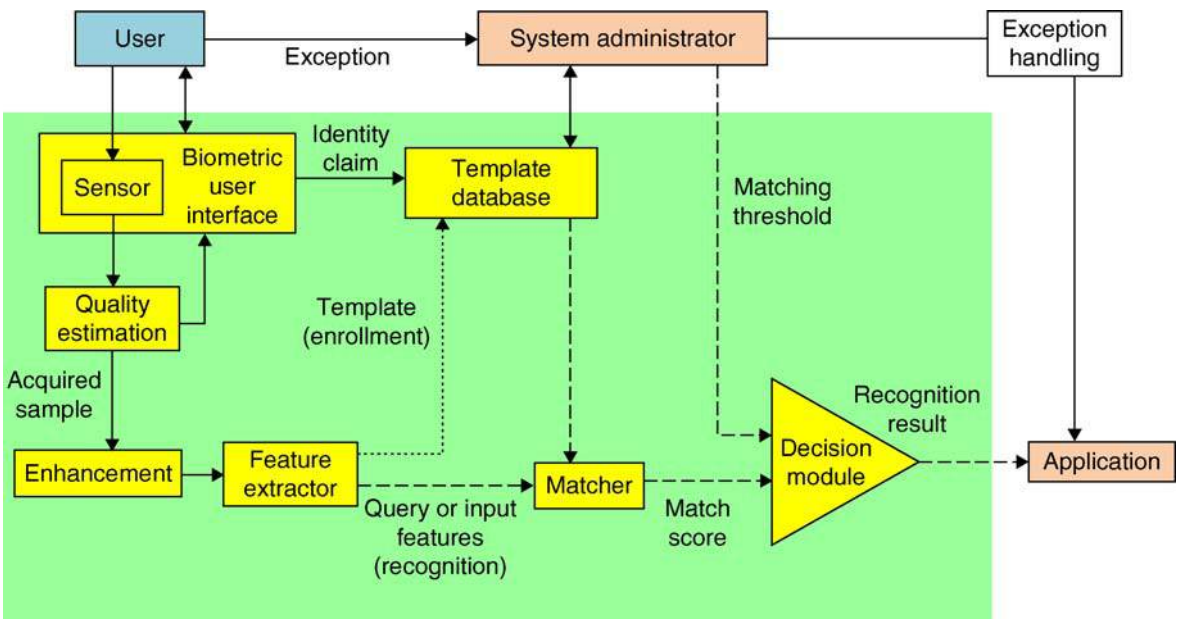
Design Specifications

The six basic design specifications [3] of a biometric system are presented below. While some of the parameters like accuracy and throughput can be measured quantitatively, factors such as security, privacy and usability are generally addressed in a qualitative manner.

Accuracy

A biometric system can make two types of errors, namely, false non-match and false match. When the intra-user variation is large, two samples of the same biometric trait of an individual (mate samples) may not be recognized as a match and this leads to a false non-match error. A false match occurs when two samples from different individuals (non-mate samples) are incorrectly recognized as a match due to large inter-user similarity. Therefore, the basic measures of the accuracy of a biometric system are *False Non-Match Rate* (FNMR) and *False Match Rate* (FMR). In the context of biometric verification, FNMR and FMR are also known as False Reject Rate (FRR) and False Accept Rate (FAR), respectively. In biometric identification, the false match and false non-match errors are measured in terms of the False Positive Identification Rate (FPIR) and False Negative Identification Rate (FNIR), respectively [4].

Accuracy requirements for a biometric system depend on the application. For example, a verification application usually involves co-operative users and may require a low FMR (0.1% or less), while a relatively high



Biometric System Design, Overview. **Figure 1** Basic building blocks of a biometric system.

FNMR (1%–5%) may be acceptable. On the other hand, a negative identification application like airport screening may require both a low FNIR to prevent undesirable individuals from circumventing the system and a low FPIR to avoid causing inconvenience (in the form of secondary screening) to the other passengers.

Throughput

Throughput refers to the number of transactions that can be handled by the biometric system per unit time. Since the input is matched only to a single template in verification, throughput is not a major issue in verification systems. However, for the sake of user convenience it is essential that the entire process of sample acquisition, feature extraction, and matching be completed within a few seconds even in verification applications. Throughput is a major concern in the identification mode because it requires matching the biometric query to all the templates in the database. Therefore, large-scale identification systems employ special schemes (both hardware and software) such as indexing, binning, or filtering to facilitate efficient searching of the database and thereby improve the system throughput.

Cost

The cost of a biometric system includes the cost of all the components of the biometric system and the recurring costs required for the operation, maintenance, and upgrade of the system. Often, there is a tradeoff between the cost of the biometric components and the performance (accuracy, throughput, and usability) of the biometric system. Furthermore, the intangible costs such as those incurred due to the errors made by the biometric system must also be considered while designing a biometric system. A thorough cost-benefit analysis is essential prior to any biometric system deployment.

Security

Since biometric systems provide a more secure and reliable authentication functionality compared to password and token-based systems, it is now being widely deployed in many real-word applications. However, the biometric system itself is vulnerable to a number of attacks [5] such as usage of spoofed traits and tampering of biometric data, communication channels, or modules. These attacks may either lead to circumvention of the biometric system or denial-of-service to legitimate users. Hence, a systematic analysis of these security threats is essential when designing a biometric system.

Privacy

While biometrics facilitates secure authentication by providing an irrefutable link to the identity of a person, it also raises privacy concerns. One major objection raised by privacy experts is the problem of function creep, where the acquired biometric data is abused for an unintended purpose. For example, allowing linkage of identity records across biometric systems may facilitate tracking of users without their knowledge. Hence, due diligence must be exercised during the design process and appropriate checks and balances must be incorporated in the biometric system to protect the privacy of users [6].

Usability

Usability of a biometric system can be measured in terms of different factors like effectiveness (Can users successfully provide high-quality biometric samples?), efficiency (Can users quickly authenticate themselves without errors?), satisfaction (Are users comfortable using the system?), and learnability (Do users get habituated to the system?) [7]. Two common metrics used to measure the effectiveness of use of a biometric system are the Failure to Enroll Rate (FTER) and Failure to Capture Rate (FTCR). If an individual cannot interact correctly with the biometric user interface or if the biometric samples of the individual are of very poor quality, the sensor or feature extractor may not be able to process these individuals. Hence, they cannot be enrolled in the biometric system and the proportion of individuals who cannot be enrolled is referred to as FTER. In some cases, a particular sample provided by the user during authentication cannot be acquired or processed reliably. This error is called failure to capture and the fraction of authentication attempts in which the biometric sample cannot be captured is denoted as FTCR. Usability depends on the choice of the biometric trait, the design of the user interface and sensor quality.

Design Issues

Given the design specifications of the biometric system and the nature of the biometric system, a system designer needs to address the following three issues systematically.

Biometric System Architecture

Architecture of a biometric system is primarily defined by the storage location of the templates and the location of the matcher. The templates (or the template database) may be stored in (1) a centralized/distibuted server, (2) local workstation at the client side, and (3) a portable device such as smart card or token that is in the possession of the user. Similarly, matching may also take place at any one of the above three sites. This allows for a wide range of possible architectures ranging from a fully centralized model, where the templates reside on the server and matching also takes place at the server, to a completely decentralized model (e.g., match-on-card or system-on-device), where all the biometric processing takes place on the device and the template never leaves the device. Other intermediate architectures are also possible. For example, the template may be stored on a smart card and during authentication, the client workstation may read the template off the card and match it with the input biometric to provide access. Note that feature extraction usually takes place only at the client side (on the local workstation or the portable device) to avoid costs involved in transmitting the raw biometric sample over a communication network.

The most important factor that decides the biometric system architecture is the mode of operation of the biometric system. While it is possible to de-centralize the database (e.g., storing the biometric templates on personalized smart cards) in the verification mode, identification mode necessarily requires centralized databases. Other characteristics of the application such as cooperative versus non-cooperative users, overt versus covert recognition, attended versus un-attended application, on-site versus remote authentication, etc. also influence the architecture of a biometric system.

In the special case of multibiometric systems [8] that involve integration of evidence from different biometric sources, the term architecture may also include the design of the fusion methodology. The fusion architecture in a multibiometric system is determined by the following three factors: (1) sources of information that need to be combined (i.e., different modalities like face, fingerprint and iris, different instances of the same trait like left and right index fingers, etc.), (2) the acquisition and processing sequence (i.e., cascade, parallel or hybrid), and (3) the type of information to be fused (i.e., features, match scores, decision, etc.).

Hardware/Software Implementation

Once the architecture of the biometric system has been defined, the system designer/integrator needs to select the appropriate hardware and software components to implement the chosen system. If the system designer also manufactures all the required components like the biometric sensor, feature extraction, and matching modules, it is relatively easy to put all these pieces together to build the complete biometric system. However, in the biometrics field, the vendors who design the biometric system or develop the application around it typically partner with another set of vendors who build the biometric hardware and software modules and create OEM (Original Equipment Manufacturer) solutions. Therefore, the following issues need to be considered by the system designers [9].

- *Sample Acquisition*: The biometric sensor or the sample acquisition hardware plays a very important role in determining the performance and usability of a biometric system. Apart from its ability to acquire or record the biometric sample of the user precisely, other factors such as the size, cost, robustness to different environmental conditions, etc. must also be considered when selecting the biometric sensor. Another problem that needs to be addressed during sample acquisition is how to deal with poor quality biometric samples.
- *User Interface*: The design of a good user interface is also critical for the successful implementation of a biometric system. An intuitive, ergonomic and easy to use interface may facilitate rapid user habituation and enable the acquisition of good quality biometric samples from the user. Demographic characteristics of the target population like age and gender and other cultural issues (e.g., some users may be averse to touching a sensor surface) must also be considered when designing the user interface.

- *Biometric Processing Components*: This includes the hardware and/or software required for performing the core biometric processing tasks of feature extraction and matching. Usually, the vendors supply software development kits (SDKs) to perform these tasks. The system designer must examine whether these components are proven and tested by reliable third party evaluations. Other factors to be considered include the cost/performance tradeoff and the availability of documentation and product support.
- *Communication Channels*: The establishment of secure communication links between the different modules of the biometric system is one of the key steps in ensuring the security of the entire system. Tamper resistance, cryptographic algorithms, and challenge-response mechanisms must be incorporated to secure the communication channels so as to avoid vulnerabilities such as denial-of-service, replay attacks, man-in-the-middle attacks, etc.
- *Database Design*: The system designer is typically entrusted with the task of storing and retrieving the biometric templates and other user information in/from a database. Therefore, the organization of the records in the database must be addressed carefully to avoid unnecessary delays that may decrease the throughput. The database design is especially important in the case of large scale identification systems.
- *Interoperability*: When a biometric system is designed using components obtained from multiple vendors, it is very important to ensure their interoperability. If possible, it is always better to use products that are compliant with the existing or emerging standards so that they can be replaced seamlessly in future. In the case of software components, the system designer must also check compliance with different operating systems and platforms.

Administration Policy

Setting the administration policy of a biometric system is one of the critical steps in ensuring the successful deployment of a biometric system. The administration policy may cover a variety of issues including:

- *Integrity of Enrollment*: The success of any biometric recognition system is mainly decided by the integrity of the enrollment process. If an adversary can enroll into the system surreptitiously (under a false identity) by producing his or her biometric traits along with false credentials (e.g., fake passports, birth certificates, etc.), the effectiveness of the biometric system gets completely nullified. Hence, the administrator needs to set appropriate policies that will guarantee the integrity of enrollment.
- *Quality of Enrollment Samples*: Enrollment is generally performed under human supervision to ensure that good quality biometric samples are obtained from the users. Furthermore, the administrator needs to define policies such as the number of enrollment samples required, the minimum sample quality required for enrollment, ways to select the best quality samples, user training, and exception procedures for persons who are unable to provide good quality samples.
- *System Configuration*: This includes setting system parameters such as the matching threshold (which determines the FMR and FNMR of the system), the number of unsuccessful trials allowed before an account is locked, the alarms to be generated, template update policies, etc.
- *Exception Handling*: Biometric systems are usually riddled with exception handling procedures (or fallback systems) to avoid inconvenience to genuine users. For example, when a user has an injury in his finger, he may still be granted access based on alternative authentication mechanisms without undergoing fingerprint recognition. Such exception processing procedures can be easily abused to circumvent a biometric system. It is very important to define appropriate policies for handling such exceptions so that an adversary cannot exploit this potential loophole easily.
- *Privacy Measures*: Given the sensitivity of the biometric information, it is essential to set policies that will prevent insiders and external adversaries from modifying or tampering the template database or using the biometric data for unintended tasks. Measures such as strict audit of access logs must be implemented to protect the user's privacy.

Summary

Designers of biometric systems need to define the system architecture, address the implementation issues,

and set the administration policies in such a way that the design specifications like accuracy, throughput, cost, security, privacy, and usability are met. However, this is generally a complicated task because some of the design requirements may be contradictory. Depending on the nature of the application, a number of tradeoffs such as cost versus accuracy, accuracy versus throughput, usability versus cost, accuracy versus security may be involved in the design of a biometric system. Optimizing these requirements so as to obtain the maximum return-on-investment is a challenging problem that requires a systematic design approach.

Related Entries

- ▶ Biometric Sample Acquisition
- ▶ Enrollment
- ▶ Interoperability
- ▶ Privacy Issues
- ▶ Security and Liveness, Overview

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An Introduction to Biometric Recognition. IEEE Trans Circ Syst Video Technol. Special Issue on Image- and Video-Based Biometrics **14**(1), 4–20 (2004)
2. Mansfield, A.J., Wayman, J.L.: Best Practices in Testing and Reporting Performance of Biometric Devices, Version 2.01. Tech. Rep. NPL Report CMSC 14/02, National Physical Laboratory (2002)
3. Bolle, R., Connell, J., Pankanti, S., Ratha, N., Senior, A.: Guide to Biometrics. Springer (2003)
4. ISO/IEC 19795-1:2006: Biometric Performance Testing and Reporting – Part 1: Principles and Framework (2006). Available at http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41447
5. Prabhakar, S., Pankanti, S., Jain, A.K.: Biometric Recognition: Security and Privacy Concerns. IEEE Security and Privacy Magazine **1**(2), 33–42 (2003)
6. NSTC Subcommittee on Biometrics: Privacy and Biometrics: Building a Conceptual Foundation (2006). Available at http://www.biometrics.gov/Documents/privacy.pdf
7. Theofanos, M., Stanton, B., Wolfson, C.A.: Usability & Biometrics: Ensuring Successful Biometric Systems (2008). Available at http://zing.ncsl.nist.gov/biousa/docs/Usability_and_Biometrics_final2.pdf
8. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. Springer (2006)
9. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition, 2nd edn. Springer (2009)

Biometric Systems, Agent-Based

Farzin Deravi
Department of Electronics, University of Kent, Canterbury, Kent, UK

Definition

Agent-based biometric systems use the computational notion of intelligent autonomous agents that assist the users and act on their behalf to develop systems that intelligently facilitate biometrics-enabled transactions, giving them the ability to learn from the users and adapt to application needs, thus enhancing recognition performance and usability.

Introduction

The ultimate effectiveness and success of biometric systems to a large part is dependent on the user experience when interacting with such systems. It is therefore essential that issues of user interaction and experience are considered when designing biometric systems. As user behavior and expectations as well as application requirements and operating conditions can vary widely, it becomes important to consider how systems can be developed that can adapt and learn to provide the best possible performance in a dynamic setting.

Here the paradigm of intelligent software agents may be effectively utilized to design and implement biometric systems that can dynamically respond to user and application needs. Intelligent autonomous agents and multiagent systems form a rapidly expanding research field [1]. Agents can be defined as software subsystems that interact with some environment and are capable of autonomous action, while representing the interests of some user or users. Such agents may know about their users' wishes and goals using a presupplied knowledge base as well as through a learning system. They can then use this knowledge to seek the accomplishment of their users' goals. While seeking such goals in a flexible response to their environment, agents may be designed to be proactive in exploiting any opportunities that may be available. They may also cooperate and compete with other agents and may have other valuable properties such as mobility and adaptability.

A group of interacting agents may be implemented to form a multiagent system (MAS) [2]. These are systems composed of multiple interacting agents that can be used to tackle applications, which are not possible to handle effectively with just a single agent and are well suited to situations where multiple perspectives of a problem-solving situation may be exploited. Interactions in a MAS may include cooperation, coordination, and negotiation between agents.

Negotiating agents are of particular importance in electronic commerce and the proliferation of Internet-based applications is a driving force for research and development of such multiagent systems [3]. Such multiagent systems when applied to user authentication applications can facilitate a bargain between the needs of the information provider for establishing sufficient trust in the user on the one hand and the confidentiality of the user's personal information and the ease of use of the system on the other hand. Such a balance may need to be achieved for each different service, transaction or session and may even be dynamically modified during use. Multiagent systems can provide an effective framework for the design and implementation of such systems.

Other areas of active research and development in the field of ▶ intelligent agents include software development environments and specialist programming and agent communication languages as well as the design of the overall architecture where layered or hybrid architectures, involving reactive, deliberative, and practical reasoning architectures continue to be of considerable interest [4].

Challenge of Complexity

The application of biometric systems in most realistic scenarios is bound to face the challenge of complexity resulting from a range of interrelated sources of variability that are likely to affect the performance and overall effectiveness of such systems.

These sources include, for example, users' physiological/behavioral characteristics, users' preferences, environmental conditions, variability of the communication channels in remote applications, and so on. If one considers the users' biometric characteristics alone, it is clear that with a widening user base it is important to consider the impact of "outliers" – those users who find it difficult or impossible to use the system. Failure to enrol on biometric systems or to consistently provide useable images for biometric matching may be due to a range of factors including physical or mental disability, age, and lack of familiarity or training in the use of the particular biometric systems deployed. In many applications, it is essential to ensure that no part of the user population is excluded from access and therefore, measures must be introduced to handle such outliers in a way that does not reduce the security or usability of the system.

One approach to address this issue, as well as to tackle the other grand challenges of biometrics such as performance, security, and privacy [5] is to adopt a multibiometric approach [6]. In multibiometric systems, information from several sources of identity are combined to produce a more reliable decision regarding identity. This may include fusing information from a number of modalities such as face, voice, and fingerprint, using a different sensor and biometric matching module for each modality. Here information may be fused at various stages of processing, including fusion of biometric features extracted from each modality (feature fusion) or fusion of matching scores after matching of each the biometric samples against the respective templates for each modality (score fusion). There is a wide, extensive, and varied literature on such multimodal identification systems [6]. While in most of the reported works, attention is generally focused on a multimodal recognition procedure based on a fixed set of biometrics, it is clearly possible to adopt a more flexible approach in choosing which modalities to integrate depending on individual user needs and constraints – thus removing, or at least reducing, the barrier to use by "outlier" individuals and facilitating universal access through biometrics.

Research has shown the potential advantages of a more flexible structure for multibiometric systems allowing an element of reevaluation and adaptation in the information fusion process [7]. Mismatched recognition and training conditions can lead to a reduced recognition accuracy when compared to matched conditions, suggesting that robust recognition may require a degree of adaptation. Inclusion of biometric sample quality information can further enhance the fusion process [8]. Here, an estimate is made of the quality of the live biometric sample and this is used to adapt the operation of the fusion module, which may have been trained earlier incorporating

knowledge from both biometric samples and their associated quality.

The move towards multibiometrics further accentuates system complexity and the burden on the biometric system users to efficiently utilize such systems. Instead of having to provide a sample for only one sensor, there is a set of sensors to interact with. There is more effort required from the user and more choices available in the design of the interaction with the system. Intelligent agents can thus provide a valuable way forward for designing and managing intelligent and adaptable user interfaces, while multiagent architectures can facilitate the negotiation of trust, security, and privacy requirements of system users.

With an agent-based biometric system selection of a variable set of biometric modalities can be accommodated to match the demands of a particular task domain or the availability of particular sensors. For example, a multimodal system should be able to deal with situations where a user may be unwilling or simply unable to provide a certain biometric sample, or where a preferred biometric modality cannot support a required degree of accuracy. The deployment of a multimodal approach, where it is possible to chose from a menu of available modalities and modes of interaction, can therefore help to overcome barriers to access. In the case of users with disabilities, where the use of a particular modality (e.g., speech) may be difficult or impossible, identity information is captured through alternative sensors to suit the user constraints.

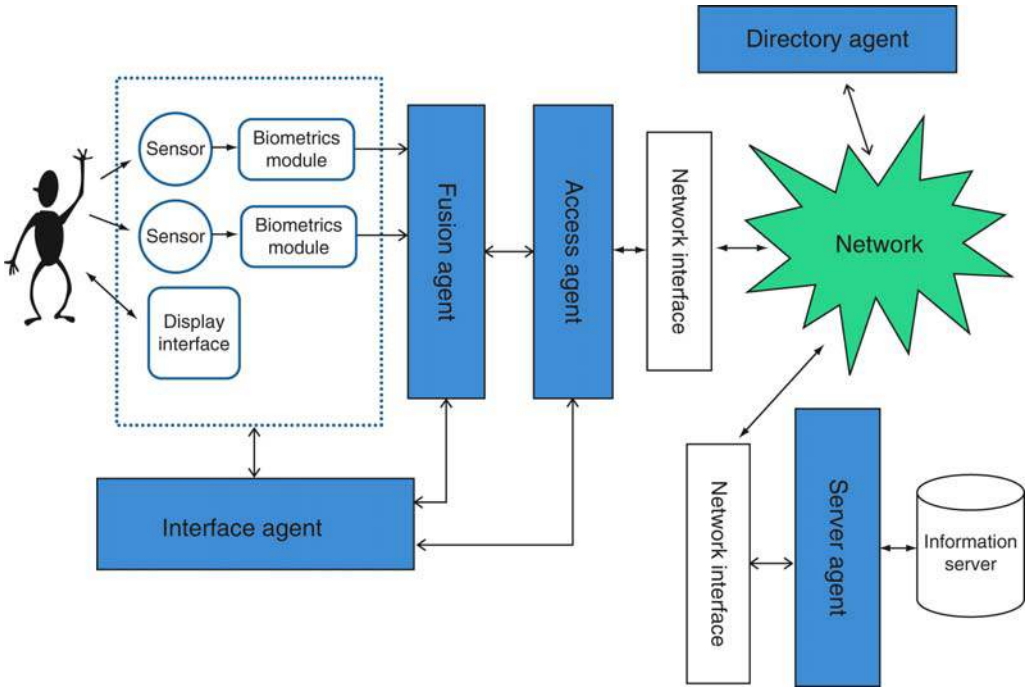
When considering remote and unsupervised biometrics-enabled access, it is essential to build in protection against attacks on the system. In particular, it is essential to establish "liveness" of the biometric input to protect against spoofing and replay attacks. This is an important consideration as for many modalities biometric samples can easily be recorded, even without the subjects' active cooperation, and it may be equally easy to present such recorded samples at the sensor or at other stages of processing to gain unauthorized access. While some progress has been made in integrating liveness detection in individual modalities, it is likely that an agent-managed multimodal framework provides a platform with additional flexibility to support more advanced robustness measures. For example, the agent interface can be deployed to provide a sophisticated challenge/response mechanism making it much more difficult to use replay attacks and much easier to establish the appropriate level of confidence in the liveness of samples.

Another important consideration when deploying biometrics in remote and networked applications is to ensure the legitimate requirements of the users at the client side to reveal only as much personal biometric information as may be necessary for establishing their access rights and no more, thus ensuring that the release of their private information is limited and controlled. At the same time, on the server side, there is the need to establish the identity of user with as high a confidence as may be required for a particular type of information access. Clearly these goals at the client and server sides are in contention and a negotiating multiagent architecture may be effectively utilized to engage in such negotiation on behalf of the users at the server and client sides.

Agent Architectures

The agent paradigm may be employed in a number of ways to enhance the performance of biometric systems. Its value is perhaps best illustrated in a multimodal biometric system for remote authentication and information access through a communication network. Here the management of the user interface, handling the information fusion process and the negotiation between the information user and information server across a network may all be delegated to a set of autonomous agents. An example of such a system for a healthcare application, using a multimodal biometric interface, has been the IAMBIC project [9], which is outlined below to illustrate possible applications of intelligent agent technologies in a biometric authentication setting (Fig. 1).

On the client side of such a client–server architecture, a set of agents will be cooperating to manage the user interface and to address the user's specific requirements and constraints. A user interface agent manages the direct interaction with the user, establishing, according to past user choices and behavior as well as the requirements of the current transaction, the set of biometric measurements that must be obtained from the user, as well as assessing the quality and reliability of the measurements from each of the available biometric recognition modules. This agent defines the mode of interaction with the user according to the user



Biometric Systems, Agent-Based. **Figure 1** A multiagent architecture for multimodal biometric authentication.

constraints and characteristics such as computer literacy, familiarity with the system being used, and so on.

The interface agent may also be responsible for the capture of other important non-biometric information. Additional environmental data may be captured by the available sensors (e.g., for the face modality a sample of background illumination may be captured). Analysis can be performed on these samples to determine the quality of any acquired data; this can be used to help the agent to analyze any possible systematic enrollment and/or verification failures. The results from this type of analysis can be used to provide feedback to the user or to system operators to improve future performance. The agent may offer immediate suggestions to a user who is finding it difficult to provide useable samples on how better to interact with the system or may request from a user whose performance has been declining over a period of time to re-enroll on to the system, thus ensuring that the biometric template ageing effects are minimized.

Additionally, the acquired samples may be associated with appropriate quality scores and this information can be passed on the fusion stage. The interface agent will also manage the individual biometric modules that will produce features and/or matching scores or decisions. Depending on the level at which the fusion takes places (sample, feature, score or decision) [10], the appropriate information is transmitted to the fusion agent to manage the fusion process.

A fusion agent can be used for the integration of the biometric measures taken from the user. Its main role is to choose the best technique for combining several different biometric measures. The design of the fusion agent requires knowledge of the types of biometrics measured, as well as of their corresponding characteristics and of the levels of confidence in claimed identity that they can typically generate. This agent may have a set of different fusion algorithms to choose from. Biometric samples, features, matching scores, or decisions obtained from the interface agent as well as sample quality and environmental information obtained from the user are passed on to the fusion agent, which in turn can produce an overall confidence score, which will be passed to the access agent for transmission to the server agent.

The access agent is then responsible for negotiating the access to the required data (e.g., medical records or other sensitive data) on behalf of the user. Essentially, this agent receives access information from the interface agent, locates the data, chooses the best location (in the event that the data can be found in different places), contacts the sources of the desired information and negotiates its release with the appropriate server agent (s). The access agent is responsible for the negotiation with the server agent and has its goal to achieve the release of the requested information. The goal of the server agent is to ensure that the information is only released to authorized users. It must ensure that sufficient confidence is reached in the identity of the claimed user. What may be considered as sufficient confidence may depend on the sensitivity of the data requested and the class of user who is accessing the information. If the result of the activities of the interface and fusion agents does not provide enough evidence to satisfy the server agent, it may enter into negotiation with them through the access agent. As part of this negotiation, a re-measurement of the biometric samples, as well as the recalculation of the combined output, may be required under specific conditions.

Optionally a directory agent may be deployed for discovering and cataloging all relevant information about location of services within the network. In a healthcare system, for instance, this agent may store information on which databases contain particular information about the patients, medical tests, and treatments. Additionally, information may be stored with regards to databases of biometric information for matching and authentication as well as information regarding, where suitable and trusted algorithms for matching, fusion, and sample quality assessment may be obtained to facilitate the agents' tasks. In the search for information, this agent may also suggest the best way of accessing required information (for example, in the situation where several databases contain the information specified), based on network traffic, distance, and so on.

Such a community of interacting agents can be implemented using a number of different methodologies for agent-based systems. These include methodologies for modeling the agents and their interactions, schemes for representing agent knowledge and languages for facilitating the communication between agents in unambiguous ways [4, 11, 12]. An important aspect of agent communication, especially in the contexts where biometrics may be involved, is to ensure the security and privacy of the information exchanged between agents. The incorporation of encryption and secure communication techniques is therefore an important consideration in the application of agent technologies to security applications.

Summary

To overcome some of the existing challenges that limit the performance and acceptability of biometric systems, as well as to develop future applications incorporating the vision of ambient intelligence, increasingly systems of greater complexity are being devised. Such systems are required to cope with large user communities, increased requirements for accuracy, security, and usability. The additional complexity of such systems provides a suitable ground for the exploitation of the intelligent agent paradigm. Multibiometric systems in particular provide a viable approach for overcoming the performance and acceptability barriers to the widespread adoption of biometric systems. An agent-based architecture can provide the support needed for the management of multimodal biometrics for person recognition and access authorization within an overall security framework for trusted and privacy-preserving information exchange.

Related Entries

- ▶ Fusion, Quality-Based
- ▶ Liveness and Anti-spoofing
- ▶ Multibiometrics
- ▶ User Acceptance
- ▶ User Interface

References

1. Wooldridge, M., Jennings, N.R.: Intelligent Agents: Theory and Practice. The Knowl. Eng. Rev. **10**(2), 115–152 (1995)
2. Jennings, N.R., Sycara, K., Wooldridge, M.: A Roadmap of Agent Research and Development. Autonomous Agents and Multi-Agent Systems, vol. 1, pp. 275–306, Kluwer, Dordecht, (1998)
3. Fatima, S.S., Wooldridge, M., Jennings, N.R.: Multi-issue negotiation with deadlines. J. Artif. Intell. Res. **27**, 381–417 (2006)
4. Weiß, G., Agent orientation in software engineering. Knowl. Eng. Rev. **16**(4), 349–373 (2001)

5. Jain, A.K., Pankanti, S., Prabhakar, S., Hong, J., Ross, A.: Biometrics: a grand challenge. Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), vol. 2, pp. 935–942 (2004)
6. Ross, A.A., Nandakumar, K., Jain, A.A.: Handbook of Multibiometrics. Springer, New York (2006)
7. Chibelushi, C.C., Deravi, F., Mason, J.S.D.: Adaptive classifier integration for robust pattern recognition. IEEE Trans. Syst. Man Cybern. B Cybern. **29**(6), 902–907 (1999)
8. Nandakumar, K., Chen, Y., Jain, A.K., Dass, S.C.: Quality-based score level fusion in multibiometric systems. Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), vol. 4, pp. 473–476. Hong Kong (2006)
9. Deravi, F., Fairhurst, M.C., Guest, R.M., Mavity, N., Canuto, A.D.M.: Intelligent agents for the management of complexity in multimodal biometrics. Int. J. Universal Access Inf. Soc. **2**(4), 293–304 (2003)
10. ISO/IEC TR 24722:2007: Information technology – Biometrics – Multimodal and other multibiometric fusion (2007)
11. Zambonelli, F., Jennings, N.R., Wooldridge, M.: Developing multiagent systems: the Gaia methodology. ACM Trans. Softw. Eng. Methodol. **12**(3), 317–370 (2003)
12. Chaib-draa, B., Dignum, F.: Trends in agent communication language. Comput. Intell. **18**(2), 89–101 (2002)

Biometric Technical Interface, Standardization

John Larmouth
University of Salford, Salford, Greater Manchester, UK

Synonyms

Biometric interchange formats; CBEFF; BioAPI; BIP; Tenprint capture

Definition

There are three main sets of international biometric technical interface standards. The first set is the Common Biometric Exchange Formats Framework (▶ CBEFF) Standards that provide for the addition of meta-data (such as date captured, expiry date, capture device information, and security information supporting integrity, and/or encryption) to a biometric data format (a fingerprint image or minutiae, an iris image, dynamic information related to a signature, etc – a Biometric Data Block, or BDB). The second set is the Biometric Application Programme Interface (▶ BioAPI) standards that provide for interchange of biometric information between modules (provided by different vendors) within a single biometric system. The third is the ▶ BioAPI Interworking Protocol (BIP) that provides for the exchange of biometric information and control of biometric devices between systems (provided by different vendors) over a network.

Introduction

This entry in the Encyclopedia describes the main standards specified by ISO/IEC JTC1/SC 37/WG2. WG2 is the Working Group responsible for Biometric Technical Interface Standards.

Biometric Data Records

There are many different forms of biometrics that can be used for human recognition (see Biometric Data Interchange Format, Standardization). These include the image of a face, a fingerprint, an iris, a signature, DNA, or a portion of speech. In general, comparison requires that features be extracted from the captured data to enable computers to identify the closeness of a match between enrolled data (data that is intended to be used for recognition purposes) and data captured for the purposes of authentication of the human being at a later time (see Biometric System Design, Overview).

There are approximately 15 standards [1] covering data interchange formats for recording such data, and all result in the specification of a Biometric Data Record – a data structure (specified down to the bit-level) that records the captured data, with different formats for the data captured before feature extraction and for that captured after feature extraction.

When used for interchange purposes with CBEFF (Common Biometric Exchange Formats Framework), a Biometric Data Record is called a Biometric Data Block (BDB), sometimes referred to as "an opaque data block".

CBEFF Wrappers

For interchange purposes, a Biometric Data Record needs to be associated with meta-data (described

below) that relates to that BDB. The package of a BDB with the meta-data (and possibly a Security Block) is then called a CBEFF Biometric Information Record, or CBEFF BIR. One of the most important pieces of meta-data is to identify (using a world-wide unambiguous identification) the BDB that is included in the BIR, without the need to know the encoding of the BDB. Without this meta-data, the nature of the BDB (finger-print, face image, etc) needs to be known by some side-channel as the BDB formats are generally not self-identifying. A point to be mentioned here is that the encodings used in current BDB formats are sufficiently similar in their initial part that intelligent software could determine which format is present, but the meta-data provides an identification without having to attempt to decode the BDB.

This is the first useful level for the interchange or storage of biometric data, unless the same modality or BDB format is used in the database or application always.

There are several forms for a BIR, designed for different applications. Some are binary-encoded, some are XML-encoded. These are described below. The format of a BIR is generally referred to as a Patron Format, as it is defined by a recognized standards development organization that is the producer of open standards – standards that are subject to vetting procedures that ensure that they are technically accurate and have wide-spread approval (a CBEFF Patron). As at 2008, there is only one registered CBEFF Patron, ISO/IEC JTC1 SC37, though others are expected to follow, and there are many registered biometric organisations.

BioAPI Interfaces and Exchanges

If a BIR has to be passed between modules from different vendors in a single system, then the interfaces between such modules need to be defined and standardized at the level of a programme language interface.

This is the purpose of the BioAPI set of standards, currently defined in terms of C interfaces, but use of other implementation languages is not precluded.

The BioAPI standard enables one or more applications to control and interact with one or more biometric devices or processes that transform a BDB (e.g., by feature extraction), typically by passing a BIR and control information in a standardized manner (allowing implementation of the relevant modules by different vendors).

BioAPI Interworking Protocol

BioAPI Interworking Protocol (BIP) is the final step in the interchange of biometric data. It builds on the BioAPI functions and parameters, but provides a bit-level specification (language and platform independent) of the protocol exchanges needed, over identified network carriers, to allow an application in one system to interact with devices in a remote system, either to control their operation and graphical user interface, or to collect a BIR (including one or more BDBs – Biometric Data Records – and security information) from them.

It is not quite true to say that BIP is the final step. There is a requirement to include in BIP transfers the transfer of certificates related to the security policy and certified security of the devices that are being used in distributed biometric capture and processing. This work is in progress in 2008, and is beyond the scope of this essay.

CBEFF

▶ Common Biometric Exchange Framework Formats, Standardization.

History and Motivation

It was recognized at an early stage that definition of formats for recording biometric data (iris, fingerprint, face, signature) etc. was not sufficient for interchange purposes, and that a minimum requirement was the addition of some meta-data. CBEFF defines the elements of such meta-data as forming a Biometric Information Record (BIR).

One important (and mandatory) element in a CBEFF BIR is to identify the format of a BDB (finger-print, face image, signature, etc.), so registration of identifiers for BDB formats (and other related formats) became an essential part of the CBEFF work.

CBEFF (Common Biometric Exchange Formats Framework) started life as a USA Standard with a

slightly different title (Common Biometric Exchange File Formats), and was proposed for fast-tracking when ISO/IEC JTC1 SC37 was first established.

In the event, it went through the normal standardization process and many changes were made during that process. CBEFF Part 1 [2] was published as an International Standard in 2006.

There are four parts to the CBEFF set of International Standards.

CBEFF Part 1 [2] defines (at the abstract level) a set of data elements that can be used to record meta-data. Note that the definition at the abstract level means that a set of values and their semantics are specified, but the multiple ways of encoding those in a bit-pattern representation are not specified at this level. Additional specifications are needed for the encoding of those values (e.g., using various forms of binary or character representation, including XML representation, and use of empty fields to denote common default values). These encoding issues are covered in CBEFF Part 3.

Some data elements are mandatory for inclusion in a ▶ CBEFF wrapper (a CBEFF Patron Format), but most are optional for use in the definition of a CBEFF Patron Format. The abstract value "NO VALUE AVAILABLE" is also frequently included for various data elements. This is important, as it enables mappings from a BIR that contains a very little meta-data to one that provides for the recording of all (meta-)data elements. The rules for this mapping are specified in CBEFF Part 1 [2]. Care should be taken when reading that a data element is "mandatory". This statement is made at the abstract level. When using an actual encoding of a header, it is always assumed that the associated patron format is known (otherwise it could not be decoded), and some patron formats can, and do, support only a single value for the "mandatory" data elements, and encode those as an empty field (zero bits, zero octets).

CBEFF Part 2 [3] (published 2006) specifies the operation of a Registration Authority that assigns world-wide unambiguous identifications for all the "things" in the CBEFF architecture that need unambiguous identifications. CBEFF Part 2 Registration is described below.

CBEFF Part 3 [4] (published 2007) defines (at the bit-level) a number of Patron formats that are of general utility (BioAPI defines another, and the profile for the sea-farer's identity card, where the encoding space is very limited). See 4.4 below.

CBEFF Part 4 [5] (work in progress in 2008) defines (at the bit-level) a Security Block format, but others are expected to be added, including a minimal one for the sea-farer's identity card. CBEFF Part 4 Security Block (SB) formats is described below.

CBEFF Part 1 Data Elements

CBEFF defines (at the abstract level – devoid of encoding) a number of data elements, with their values, and the semantics of each value.

It also defines an architecture, where there is normally an SBH (Standard Biometric Header) that contains the meta-data elements, a BDB, and (optionally) a Security Block (SB) that contains details of encryption and integrity parameters. This is depicted in Fig. 1.

The following summarizes the data elements (meta-data) currently defined in CBEFF Part 1.

CBEFF Version: The version of the CBEFF specification used for the elements of the SBH.

BDB Format owner and format type: These meta-data elements identify the (registered) biometric organization that has defined the BDB format and the identifier (typically an integer from zero upwards) that has been registered as its identification (see CBEFF Part 2 Registration). They are mandatory in a BIR, and identify the BDB that is contained in the BIR. A point to be mentioned here is that there are BIR formats that contain multiple BDBs, but discussion of these is outside the scope of this essay.

BBD Encryption and BIR integrity options: These meta-data elements are mandatory, but are simple binary values saying whether the BDB is encrypted or not, or whether there is an integrity value for the BIR provided in a Security Block. If either of these is "YES", then the Security Block has to be present to provide the necessary security details, otherwise the Security Block is absent. We are operating here at the abstract level. A particular patron format may support

Biometric Technical Interface, Standardization. Figure 1 A simple BIR.

only one of these values. If only one is supported by a particular patron format (e.g., NO encryption, NO integrity), then these values can be encoded as a null coding (depending on the nature of the encoding), so need not take up bit-space (which matters for some applications).

BDB Biometric type and sub-type: This provides a broad identification of the nature of the BDB. Its value can be deduced from the "Format owner and format type", but only through the registration authority, and it is not computer friendly. It identifies the broad nature of the format (finger, face, signature, ear, iris, vein, etc.), with the subtype identifying which finger or ear or iris, etc. The categorization is a bit ad hoc, and has changed over time, and will probably continue to change.

BDB Product owner and product type: These two data elements identify the owner (a registered biometric organization – see CBEFF Part 2 Registration) and identification of the device/software used to produce the BDB.

BDB Creation date and validity period: The date on which the BDB was created, and the start and end of validity period. The use of the validity period is depends on the application.

BDB Purpose: This identifies the reason for the capture of the BDB – for enrolment or for verification (and there are other options). The use of this field in actual applications is not clear yet.

BDB Processed level: Again, this is implicit in the registered identifier, but it gives a broad indication of whether this is "raw" data, an enhanced image, or a format that has extracted features from an image. Values are "raw", "intermediate", or "processed", which are very broad terms. The author is not aware of systems that use or require this information.

BDB Quality: This is quite an important field, but there is still a lot of work ongoing to determine "quality" values for a BDB. It relates to whether a fingerprint is known to be smudged or not, how many pixels were used in the capturing of an image, whether a signature had enough turning points for minutiae extraction, etc. Work is ongoing in this area. (See Biometric Sample Quality, Standardization). It is likely that when the ongoing work is completed, this part of the Standard will be amended.

BDB Index: A meta-data element that can be used to point to a database entry for the BDB, rather than having the BDB encoded as part of the BIR. The use of this for storage is clear, but it is arguable that it is not needed, as the BIR is only defined at the abstract level, so encoding a BDB is not needed. The author is not aware of any current use.

Challenge/response: This provides data for security purposes when trying to retrieve the associated BDB from a database (like the registration procedure followed in a bank where a question is asked (e.g., "a favourite book") and the response to that). It is not yet clear as to how this field can be practically used.

Security Block (SB) Format owner and type: These meta-data elements identify the (registered) biometric organization that has defined the SB format, and the identifier (typically an integer from zero upwards) that has been registered as its identification (see CBEFF Part 2 Registration). They are mandatory if a security block is included.

BIR Creator, creation date, and validity period: These data elements recognize that the BDB may have been created at a certain time, but that this BIR (following possible processing – perhaps on a remote machine) may have been produced by a different vendor at a different time. The "creator" is just a string of Unicode characters, is not registered, and hence, may not be unambiguous. Examples of a "creator" might be "US Dept of State" or "Passport Australia".

BIR Patron Format Owner and type: The main (probably the only) use is in the complex BIR format, when a different BIR can be embedded in a simple BIR, and BIR Patron Format Owner and type identifies the nature and encoding of the embedded BIR.

BIR Patron header version: A version number (major and minor) assigned in the patron format definition.

BIR Index: A self-reference to a database entry for this BIR. The author is not aware of its any current use.

BIR Payload: A transparent string of octets associated with the BDB. The author is not aware of its any current use.

Sub-header count: This is a device to handle a BIR that contains multiple BDBs with different SBHs applied to each. The details are out of the scope of this essay.

CBEFF Part 2 Registration

The CBEFF Part 2 Registration provides for the worldwide unambiguous identification of:

- Biometric Organizations and Biometric Patrons
- Biometric Data Block Formats (BDB formats)
- Patron formats (specific selections of meta-data, with a bit-level encoding)
- Security Block formats
- Biometric products (devices and/or software modules)

  The identification is composed of three components:
  - Arcs of the International Object Identifier tree that identify the register (implicit)
  - A registered 16-bit identifier that identifies a biometric organization (of which the biometric patrons are a subset)
  - An identification assigned by the biometric organization to a BDB format, a Patron Format, a Security Block format, or a biometric product.

The CBEFF register is currently (2008) maintained by the International Biometric Industry Association (IBIA), and is available at URL http://www.ibia.org/cbeffregistration.asp.

There are a large number of biometric organizations registered, a few products, but in 2008 only ISO/IEC JTC 1 SC 37 has registered BDB formats, Patron Formats, or Security Block formats.

CBEFF Part 3 Patron Formats

The CBEFF Part 3 Patron Formats specifies a range of Patron Formats designed for use in the areas of different application. The smallest is the minimal binary encoding, where most elements take only a fixed value (typically "NO VALUE AVAILABLE" if the element is optional), and produce zero bits in the encoding. There are other formats that produce XML encodings for the data elements, and are capable of encoding the complete range of abstract values of every element.

Some Patron Formats are defined in English with a tabular format for the bit-patterns, so no tool support is available for these.

Others are defined using the ▶ Abstract Syntax Notation One (ASN.1), the notation [6] which (provides a schema for both binary and XML encodings) is defined using both XSD (XML Schema Definition [7]) and an equivalent ASN.1 schema for an XML encoding, in addition to the English language specification.

Both the ASN.1 and XSD schemas are supported by a range of tools on many platforms.

In 2008 there are 17 Patron Formats defined and they are:

- Minimum bit-oriented: This takes only one octet for the SBH if the BDB format owner is SC 37 and the format type value is less than 64. It is default in all fields to fixed values apart from the BDB format owner and format type. The specification uses the ASN.1 notation and the ASN.1 Unaligned Packed Encoding Rules.
- Minimum byte-oriented: This takes four octets and is specified with tables, diagrams, and English language.
- Fixed-length fields, byte-oriented: This can handle all data elements (with some length restrictions), but optional ones that are absent (NO VALUE AVAILABLE) encode with a single "presence" bit of zero. The specification uses tables and English language.
- Fixed-length fields, bit-oriented: This can handle all data elements, of arbitrary length (so length fields are frequently present), but optional ones that are absent (NO VALUE AVAILABLE) encode with a single "presence" bit of zero. The specification uses ASN.1 and the ASN.1 Unaligned Packed Encoding Rules.
- Full support, TLV format: This can handle all data elements. Length fields are always present, and every element is preceded by an identifying tag (or type) field. It is based on the earlier use in smart-cards, and uses an ASN.1 specification with the Type Length Value (TLV) Basic ASN.1 Encoding Rules.
- This supports nested BIRs within BIRs: Specified with tables and English, with supporting ASN.1.
- XML encoding: Specified with tables and English language, with supporting ASN.1 (XML Encoding Rules) and XSD specifications.

There is also a Patron Format defined in BioAPI, largely for historical reasons.

CBEFF Part 4 Security Block (SB) Formats

CBEFF Part 4 Security Block (SB) Formats is in progress in 2008, so a detailed discussion is not appropriate. At present there is only one Security Block format being defined, that handles all necessary security

parameters for either encryption or integrity, or both, and allows the use of a wide range of security algorithms.

It is likely that a more minimum SB format will be defined for use with the seafarers' identity card (a standard being progressed by ISO/IEC JTC1 SC37), and handles only integrity with fixed algorithms (See Biometrics Security, Standardization.).

BioAPI

► Biometrics API/Interfaces

History and Motivation

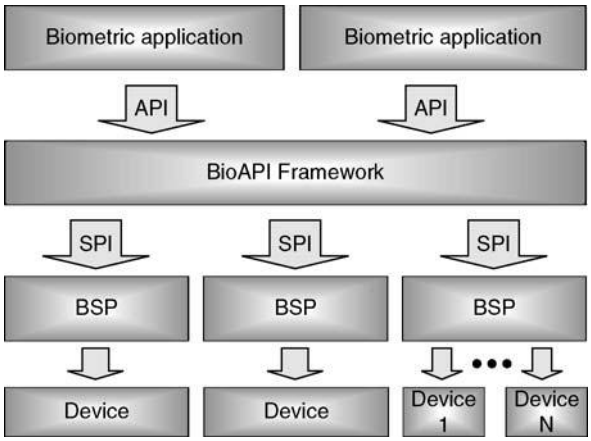
Multiple application modules (perhaps from different vendors) should be allowed to interact (serially or simultaneously) with multiple modules supporting various biometric devices. Standard interfaces are needed to allow these modules to potentially be provided by different vendors.

The concept of a "framework" module, with which applications attach above and device-related software attaches below, providing a general routing function for commands and data transfer, is the main part of the BioAPI architecture.

There are four groups of standards in the BioAPI set.

The first is the base standard – BioAPI Part 1 [8] (published in 2006, but with several amendments to extend its functionality). This part defines the concept of the ► BioAPI framework module which interacts above with applications, using a C-defined API, and below with Biometric Service Providers (software and hardware related to biometric devices) using a C-defined Service Provider Interface (SPI), broadly mirroring the functionality of the API. This is illustrated in Figure 2 – BioAPI architecture. It also has a specification for Graphical User Interface to enable an application to control the "screens" for use during a capture operation.

The second group (currently only BioAPI Part 2 [9] and Part 4 [10]) is a set of standards providing a lower-level interface within a BSP to a so-called "function provider module" that is distinct from the vendor of the BSP module. This interface is designed to minimize the requirements on a device vendor, and to enable the provider of software for a BSP to use modules from many different device vendors. Detailed interfaces are not covered in this article.



Biometric Technical Interface, Standardization.
Figure 2 BioAPI architecture.

Part 2 [9] was published in 2007 and provides an interface to archive devices (databases). Part 4 [10] is in progress in 2008, and provides an interface to sensor (capture) devices. Similar interfaces for matching algorithm modules and general processing modules are planned, but have not been started in 2008.

The third is a BioAPILite standard (BioAPI Part 3 [11]) that is intended to provide support for embedded devices. This is not mature in 2008, and will not be discussed further.

The fourth is a standard [12] specifying how to use the BioAPI interfaces to support the so-called "ten-print slap" – a roll of ten fingers, an image of four left fingers, an image of four right fingers, and an image of two thumbs – and the subsequent processing of the returned BDB, possibly to extract parts of the BDB to individual BDBs.

In fact, this standard is rather more general than just supporting a ten-print slap, and recognizes the concept of gathering data into a single BIR from a number of different biometric devices, possibly of different modalities. This introduces a new BIR concept of a complete (complex) BIR with "holes" in it (place-holders) that will be filled in whenever possible when passed to a BSP, and can then be passed to other BSPs to complete it. The interesting thing is that this development does not require any change to the basic BioAPI architecture or function calls – these already allowed the transfer of a BIR to a BSP (e.g., for image enhancement purposes),

with return of a new BIR. It is in progress in 2008, and is not discussed further, but is likely to become important.

BioAPI Part 1

The two interfaces (API and SPI) are very similar, as the framework provides mainly routing and (when augmented with BIP functionality – see Biometric Interworking Protocol (BIP) below) communications functionality to remote systems.

Indeed, there is an amendment to Part 1 that is being developed which recognizes the use of a reduced API/SPI to provide a direct interface between an application and a BSP, with support for multiple BSPs, or multiple applications being done entirely through the (non-standardized) operating system. This is called "frameworkless BioAPI."

BioAPI generally assumes that the BSP is not state-free, so there can be a request for a BDB to be captured, and a "handle" returned pointing to it. It is stored in memory controlled by the BSP, and later "exported" to the application through a subsequent call. Thus, there are several memory management functions and parameters.

The normal sequence of interaction between any application module and (through the framework) a BSP module is described below. Note that there can be multiple such simultaneous interactions related to one application and multiple BSPs or BSP instances, or related to one BSP and multiple application modules or instances. The normal sequence has some options within it (controlled by the application), and there can be a variety of error returns or signals that can disrupt the sequence. There are a variety of parameters that can be passed by the application to control the way the BSP operates, but these are beyond the scope of this article. The normal sequence is:

- *Init:* This introduces the application instance module to the framework module, and establishes that they both are using the same version of the interface specification.
- *Load BSP:* This tells the framework that (at least one) application instance wants to communicate with it.
- *BSP attach:* This initiates a dialogue with the BSP, and establishes an error reporting mechanism.

  Enroll for verification: This initiates a capture, and returns a BDB, suitable for enrolment of the subject; or

  Verify: This initiates a capture, and returns a BDB, suitable for verification against a previously stored biometric reference or template.
- *BSP close:* This says that the application is no longer interested in interactions with the BSP.

Of course, multiple calls between attach and close are possible. There are also calls to the framework to establish what BSPs are available, and their properties, but this is too detailed for this essay.

Tool Support

Implementations of the framework module are available from a number of vendors.

Implementations of BSPs that support the standardized (SPI) interface are still emerging (2008), as are application modules using the BioAPI API interface.

Biometric Interworking Protocol (BIP)

History and Motivation

The need for an application to interact with remote biometric devices (or with modules processing and transforming biometric data) over a network, in a fully standardised manner (providing vendor independence of the communicating systems) in the standardization process was recognized early.

BioAPI was seen as the appropriate base for this. Essentially, the BIP specification extends the functionality of a BioAPI framework to allow it to route calls from an application to a remote framework (and hence a remote BSP) and to support the return of appropriate results.

It also supports the provision of a remote Graphical User Interface (screen), controlled by a remote application, to perform a capture.

Fundamentally, it provides a mapping from the BioAPI Part 1 C-functions and data structures into protocol elements and ASN.1 data structures that are then encoded with the ASN.1 unaligned Packed Encoding Rules.

This means that a BIP-enabled framework can communicate with another BIP-enabled framework for communication between local applications and remote BSPs (or vice versa).

It is important to note that a computer system can support BIP if it provides the appropriate "bits-on-the-line" exchanges that would occur if it had a BioAPI framework module. The BIP specification is based on BioAPI, but is a fully-defined protocol that creates no constraints on the internal structure of the communicating systems. In terms of communication "bits-on-the-line", internal module structure is invisible and irrelevant. The concept of a BioAPI framework is used in the specification of the messages, but that does not need to form a part of the internal structure of the communicating systems.

Supported Network Mappings

The BIP Standard is a fully defined protocol over TCP/IP (the Internet) using a recommended port of 2376, registered with the Internet Assigned Numbers Authority (IANA).

It also specifies discovery and announcement protocols based on both IPv4 and IPv6. It also specifies its use over W3C SOAP/HTTP.

Tool Support

There are many tools supporting ASN.1 defined protocols that can be used, but there are some vendors already advertizing full BIP support within a BioAPI Framework.

Related Entries

- ▶ Biometric API/Interfaces
- ▶ Biometric Data Interchange Format, Standardization
- ▶ Biometric Sample Quality, Standardization
- ▶ Biometric System Design, Overview
- ▶ Biometric Vocabulary, Standardization
- ▶ Biometrics Security, Standardization
- ▶ Common Biometric Exchange Framework Formats, Standardization
- ▶ International Standardization of Biometrics, Overview
- ▶ MultiBiometrics and Data Fusion, Standardization

References

1. All parts of ISO/IEC 19794 *Biometric Data Interchange Formats*
2. CBEFF Part 1 (ISO/IEC 19785–1) *Data Element Specification*
3. CBEFF Part 2 (ISO/IEC 19785–2) *Procedures for the Operation of the Biometrics Registration Authority*
4. CBEFF Part 3 (ISO/IEC 19785–1) *Patron formats*
5. CBEFF Part 4 (ISO/IEC 19785–1) *Security Blocks*
6. ASN.1 (ISO/IEC 8824–1) *Abstract Syntax Notation One*
7. XSD *W3C XML Schema*
8. BioAPI Part 1 (ISO/IEC 19784–1) *BioAPI Specification*
9. BioAPI Part 2 (ISO/IEC 19784–2) *Archive Function Provider Interface*
10. BioAPI Part 3 (ISO/IEC 19784–3) *BioAPILite*
11. BioAPI Part 4 (ISO/IEC 19784–4) *Function Provider Interface*
12. BioAPI Ten-print (ISO/IEC 29129) *Tenprint capture using BioAPI*

Biometric Technology Test

▶ Performance Evaluation, Overview

Biometric Template

A biometric template is a digital representation of an individual's distinct characteristics, computed or extracted from a biometric sample. It is biometric templates that are actually compared in a biometric recognition system. The forms of biometric templates can vary between biometric modalities as well as vendors. Not all biometric devices are template based. For example, voice recognition is based on speaker models.

- ▶ Biometrics, Overview
- ▶ Identification
- ▶ On-Card Matching
- ▶ Verification

Biometric Terminal

A terminal, which comprises of processing element (PC or embedded system), biometrics sensor, card reader, and optional network access; captures a presented

biometric trait, for example, face or fingerprint; and is able to encode the captured biometric data into a template for identity verification is biometric terminal.

► On-Card Matching

Biometric Transaction Time

► Operational Times

Biometric Variability

Biometric variability refers to the differences in the observed features from one instance of the biometric to another. The differences can be random, or systematic due to some underlying factor that governs the variability.

► Individuality of Fingerprints: Models and Methods

Biometric Verification/ Identification/Authentication/ Recognition: The Terminology

James L. Wayman
College of Engineering, San Jose State University, San Jose, CA, USA

Synonyms

Biometric authentication; Biometric identification; Biometric verification

Introduction

There has been an inconsistency in the use of the terms like "recognition," "authentication," "identification," and "verification," throughout the literature of biometrics. Particularly, with the applications of automated human recognition technologies becoming more creative, older uses of these terms has become inadequate in describing new systems. In this article, the author will explore some of the historical uses of these terms and suggest some definitions consistent with recent applications of the technologies.

Dictionary Definitions

The essay begins with common, natural language definitions of these four terms. The Oxford English Dictionary [1] provides definitions for the terms discussed in this article:

authenticate: prove or show to be authentic
authentic: of undisputed origin or veracity; genuine
recognition: the action or process of recognizing or being recognized
recognize: identify as already known; know again
verification: process of verifying; the establishment by empirical means of the validity of a proposition
verify: make sure or demonstrate that (something) is true, accurate or justified.
identification: the action or process of identifying or the fact of being identified
identify: establish the identity of.
identity: the fact of being who or what a person or thing is

Historical Usages

Since the earliest literature of biometrics, a difference in functionality of automated human recognition technologies has been discussed. In 1966, Li et al. [2] wrote:

> To simplify this study, the problem was confined to the verification (or rejection) of an utterance as that of an expected informant. This process is defined as *speaker verification* (as opposed to *speaker identification*, which is the selection of an actual speaker from a population) (Italics in the original)

Verification and identification are defined here as two, mutually exclusive applications for biometrics – verification as the recognition of an expected person and identification as the selection of a person out of a population. The 1960s biometrics literature, however, was far from consistent in the use of these terms [3–5].

In 1969, IBM [6] listed "four principal differences that distinguish a verification procedure from an identification procedure" as:

1. An alien class, for which a priori information is not available, is considered by the system.
2. Additional information, the class label, is available for the decision.
3. The class label that is entered can determine which parameters are to be extracted from the pattern.
4. The decision involves only two states, acceptance or rejection of the pattern.

In identification, all possible classes are presumed known, and the decision amounts to the best match of the pattern to a particular class.

The above quote uses the word "class" to mean a "person." The IBM definitions attempt to be more precise, but limit "identification" to the case where all possible persons ("classes") are known, though the more common case in practice accepts that previously and unknown persons can be encountered. Tosi [7] and Tosi, et al. [8] differentiated identification into "closed trials" (all persons known) and "open trials" (unknown persons presumed to exist). Modern parlance calls these as "open-set" and "closed-set" identification. In closed-set identification, the question asked is "If any, which of the known persons are consistent with the encountered data?" and "None" is an appropriate response in open-set identification, but not a possible response with a closed-set. Closed-set identification is the easier task, as the person represented by the data is guaranteed to be among those known to the system.

Real-world applications are almost always open-set [9], allowing for the possibility of encountering someone who is not enrolled (an "impostor"). Some academic communities within the field of biometrics, however, currently define identification solely as "closed-set." This community reports, as the outcome of the identification task, an ordering of the enrolled biometric data by similarity to the submitted sample. A sample is considered to be "recognized" if it is among the highest k members of the list, where k is determined by the researcher [10–14].

US government standards in the 1970s [15] did not differentiate between open- and closed-set identification, but differentiated between "'absolute' identification" and "verification of identification" – the former what is now called as "identification" and the latter as "verification."

In the 1980s, the International Biometrics Association (IBA) attempted to create a standard set of definitions for use in biometrics [16]. This vocabulary defined verification saying, "Verification of identity is the operation of comparing a submitted biometric sample against a specific claimed biometric reference template to determine whether it sufficiently matches that template." The IBA did not attempt a definition of "identification," but offers the definition of "recognition" as: "Recognition of identity is the operation of comparing a submitted biometric sample against the population of biometric reference templates to determine whether it belongs to the population and which member of the population it is." This definition seems consistent with what has been called open-set identification in this essay.

The classical "verification" concept, as defined in [2] quoted above, can be implemented with either a centralized database of stored references for each enrolled user, or with a tamper-proof token, such as a passport or card, that carries the reference. For centralized systems, data subjects (the users of the system) must point to their stored reference in some way – either with a PIN, a card, or a unique name. This pointer must refer to only the references of a single enrolled user. Therefore, data subjects cannot be free to choose their own identifiers, but each must be assigned a different identifier.

The Impact of New Algorithms on Terminology

By the mid-1990s, vendors had introduced the term "PIN-less verification" to denote access control systems that did not require data subjects to submit a user identifier with their biometric sample. These systems had the internal programming of an "identification system," examining all enrolled references to determine if the submitted sample was similar to any, but had the external look and feel of a "verification" system. By the mid-1990s, all commercially-available iris recognition access control systems were based on the "PIN-less verification" concept, allowing users physical and logical access without submission of a user identifier. Iris systems modified to allow the input

of a user PIN would still search the entire database for a matching reference iris pattern, and then compare the PIN stored with the matched pattern to the PIN submitted to further validate the match.

Other approaches in the 1990s [17], based on the concept from forensic fingerprinting of "binning" all similar fingerprints together, allowed users to choose their own identifying PIN or password (which would denote the "bin") with the understanding that many users might chose the same one. The system would have to compare the submitted biometric samples against the stored references of all users within the "bin" denoted by the submitted PIN. As users would not know how many other references, if any, were identified by the same PIN, such systems would again have the look and feel of a "verification" system, though performing an open-set "identification" function against a group of users. In the case if there was only one user with a particular PIN, the system would degenerate into "verification" as defined in [2]. In other words, the system was performing either "verification" or open-set "identification" depending upon the number of users stored with the PIN that was entered.

By the end of the 1990s, it was clear that there were no longer clear boundaries between "verification" and "identification," the differences depending upon the specifics of the algorithm and the stored data. Using the classic definitions, many applications could not be clearly determined as "verification" or "identification."

Clarifying Meanings

By the early 2000s, clarification of this confusion was clearly needed so that an application could be described independently of the details of the algorithms and data structures used to instantiate it. A study by the U.S. National Research Council (NRC) sought to restore the usability of the terms "authentication" and "identification" [18] in discussion of general computer-based methods for determining user credentials. This study returned to the dictionary definitions above, defining "authentication" as "the process of establishing confidence in the truth of some claim" and "identification" as "the process of using claimed or observed attributes of an individual to infer who the individual is." The term "verification" was considered

Box 1: Current International Definitions from ISO/IEC JTC1

biometric ***verification*** (biometric application)

application that shows true or false a claim about the similarity of biometric reference(s) and **recognition** biometric sample(s) by making a comparison(s)

EXAMPLE: Establishing the truth of any of the claims "I am enrolled as subject X", "I am enrolled in the database as an administrator", "I am not enrolled in the database", may be considered verification.

NOTE: A claim of enrolment in a database without declaring a specific biometric reference identifier may be verified by exhaustive search.

closed-set ***identification*** (biometric application)

application that ranks the biometric references in the enrolment database in order of decreasing similarity against a **recognition** biometric sample

NOTE 1 Closed-set **identification** always returns a non-empty candidate list

NOTE 2 Closed-set **identification** is rarely used within practical systems, but is often used experimentally

open-set ***identification*** (biometric application)

application that determines a possibly empty candidate list by collecting one or more biometric samples from a biometric capture subject and searching the enrolment database for similar biometric references

NOTE Biometric references may be judged to be similar on the basis of comparison score.

authentication

NOTE 1 Use of this term as a synonym for "biometric **verification** or biometric **identification**" is deprecated; the term biometric recognition is preferred.

by the NRC committee as a term used primarily within the biometrics community and synonymous with "authentication."

Since at least the 1990s [19], it had been noted that biometric claims could be negative as well as positive – for example, "I am not data subject X" or "I am not enrolled in the biometric system." Using the NRC definitions, the term "authentication" could be used to describe the process of establishing the truth of such a negative claim and "identification" could be used to describe the outcome of an access control system.

Applying these definitions leads to some clarity of language, restoring the dictionary, natural-language meanings of these terms. "Verification" and "authentication" can apply to positive or negative claims. A data subject, or some other party, need not specify an identifier, such as a PIN, pointing to an enrolled biometric reference. So, for example, biometrics can be used without a user identifier to verify that I am enrolled in the system or that I am not enrolled in the system. Examples of the former are biometric systems used to prevent issuance of multiple enrolment records to the same user. Examples of the latter are often called "watchlists." With a claimed user identifier, biometric systems can verify that I am enrolled (known to the system) as X, or not enrolled as X. A consequence of this definition is that all biometric systems can be seen as verifying some kind of a claim, whether positive or negative, whether with or without a specified user identifier. The details of either the algorithm or the data structures need not be considered in applying the term "verification." "PIN-less verification" systems are indeed "verification" systems.

Under the NRC definitions, "identification" is the process of "infer(ring) who the person is," meaning to return an identifier (not necessarily a name) for that person. This process can include a claim to an identifier by the data subject or by someone else in reference to the data subject (i.e., "She is enrolled as user X"). By these definitions, "identification" and "verification" are not mutually exclusive. A biometric system can identify a person by verifying a claim to a known identity. This usage is consistent with the historical documents such as [4, 5].

At the time of writing this essay, the international standards committee on biometrics, ISO/IEC JTC1 SC37, has tentative definitions for the terms considered in this article [20], shown in the box below. The SC37 definitions are compatible with those of the NRC, although SC37 prefers "biometric verification" to "biometric authentication," the latter being depreciated in the vocabulary corpus. The SC37 definitions do not include "recognition," deferring to common dictionary definitions for the meaning of that term.

Related Entries

▶ Biometrics, Overview

References

1. Pearsall, J. (ed.): Concise Oxford English Dictionary, 10th revised edn. Oxford University Press (2002)
2. Li, K.P., Dammann, J.E., Chapman, W.D.: Experimental studies in speaker verification using an adaptive system. JASA. **40**(5), 966–978 (1966)
3. Pruzansky, S.: Pattern matching procedure for automatic talker recognition. JASA. **35**(3), 345–358 (1963)
4. Mauceri, A.J.: Technical Documentary Report for Feasibility Study of Personnel Identification by Signature Verification, North American Aviation, Inc, Space and Information Systems Division, SID 65–24, accession No. 00464–65, 19 January, 1965
5. International Business Machines Corporation.: The Considerations of Data Security in a Computer Environment, IBM Corporation, White Plains, NY, Form 520–2169, 1969
6. Dixon, R.C., Boudreau, P.E.: Mathematical Model for Pattern Verification, IBM J. Res. Dev. November, 717–729, 1969. Available at http://www.research.ibm.com/journal/rd/136/ibmrd1306I.pdf
7. Tosi, O., Oyer, H., Lashbrook, W., Pedrey, C., Nicol, J., Nash, E.: Experiment on voice identification. JASA. **51**(6), 2030–2043 (1972)
8. Tosi, O.: Experimental Studies on the reliability of the voiceprint identification technique. In: Proceedings of third National Symposium of Law Enforcement and Technology, Chicago, IL (1970)
9. Doddington, G.: Speaker recognition evaluation methodology: a review and perspective. In RLA2C, Avignon, April 1998, pp. 60–66
10. Li, F., Wechsler, H.: Robust part-based face recognition using boosting and transduction. In: IEEE Conference on Biometrics: Theory, Applications and Systems, Washington, DC, 27–29 September 2007
11. Bolme, D.S., Beveridge, J.R., Howe, A.E.: Person identification using text and image data. In IEEE Conference on Biometrics: Theory, Applications and Systems, Washington, DC, 27–29 September 2007

12. Gross, R.Y.L., Sweeney, L., Jiang, X.Q., Xu, W.H., Yurovsky, D.: Robust hand geometry measurements for person identification using active appearance models. In: IEEE Conference on Biometrics: Theory, Applications and Systems, Washington, DC, 27–29 September 2007
13. Arbab-Zavar, B., Nixon, M.S., Hurley, D.J.: On model-based analysis of ear biometrics. In: IEEE Conference on Biometrics: Theory, Applications and Systems, Washington, DC, 27–29 September 2007
14. Cadavid, S., Abdel-Mottaleb, M.: Human Identification Based on 3D Ear Model. In: IEEE Conference on Biometrics: Theory, Applications and Systems, Washington, DC, 27–29 September 2007
15. National Bureau of Standards.: Guidelines for evaluation of techniques for automated personal identification. Federal Information Processing Standards Publication 48, April (1977)
16. International Biometrics Association(IBA): "Standard Biometric Industry Terminology and Definitions: Draft", IBA Standard BSC 2.6 – 1987R, Washington, DC, Oct. 27, 1987
17. Pare, D.F., Jr., Hoffman, N., Lee, J. A.: Tokenless identification of individuals. US Patent 5,805,719, 8 Sep, 1998
18. Kent, S., Millett, L.: Who goes there? Authentication technologies through the lens of privacy National Academies Press, Washington, DC, 2003. Available at http://www7.nationalacademies.org/cstb/pub_authentication.html
19. Wayman, J.L.: Fundamentals of biometric authentication technology. Proceedings CardTech/Securtech, Chicago, IL, May 11–14, 1999. Available at http://www.engr.sjsu.edu/biometrics/publications_fhwa.html
20. Standing Document 2 – Harmonized Biometric Vocabulary, version 7, ISO/IEC JTC 1/SC 37N1978, 12 February 2007

Biometric Vocabulary, Standardization

Rene McIver
Keewatin Avenue, Toronto, ON

Definition

The International Standards Organisation (ISO) Standard 1087-1 [1] defines *vocabulary* as a "*terminological dictionary which contains a list of designations and definitions from one or more specific subject fields*". For the subject field of *biometrics*, while there are several publications of biometric vocabularies, including [2, 3], there is no single collection of biometric terms and definitions considered by the industry to be the definitive source. As a result, there are inconsistencies across biometric literature which may negatively impact knowledge representation and transfer. Efforts are underway in ISO/International Electrotechnical Commission (IEC) Joint Technical Committee (JTC) 1 SC 37, however, to harmonize the biometric terminology that exists in the industry and develop a biometric vocabulary that will surely become the definitive source.

Terminology Development in ISO

Three ISO standards [1, 4, 5] are currently in publication that provide guidance for terminology work useful both inside and outside the framework of standardization. ISO 1087-1 [1] consists of a set of terms and definitions required for terminology development.

The following are of particular importance

Characteristics: Abstraction of a property of an object or of a set of objects
Concept: Unit of knowledge created by a unique combination of *characteristics*
Concept diagram: Graphic representation of a *concept system*
Concept system: Set of *concepts* structured according to the relations among them
Definition: Representation of a *concept* by a descriptive statement which serves to differentiate it from related concepts
Subject field: Field of special knowledge
Term: General designation of a general *concept* in a specific *subject field*

With the vocabulary needed for terminology work is provided in [1], ISO 704 [4] establishes a common framework of thinking to be used, when developing a terminology. It outlines links between objects, concepts, and their representations, as well as general principles in term and definition formulation.

To state simply, terminology development involves:

- Identifying concepts and understanding their characteristics
- Grouping of related concepts into concept systems
- Representing concept systems through concept diagrams

- Defining concepts
- Attributing terms to each concept

It is important to understand the characteristics of each related concept on a concept diagram to ensure each is truly distinct and all concepts have been identified. Once this is accomplished, definition crafting becomes a simple matter of wordsmithing using only those characteristics deemed essential for the concept.

ISO 860 [5] serves to close a potential gap in terminology development through *concept* and *term harmonization*:

Concept harmonization: Activity for reducing or eliminating minor differences between two or more concepts that are already closely related to each other

Term harmonization: Activity relating to the designation of one concept in different languages by terms that reflect the same or similar characteristics or have the same or slightly different forms

Concept and term harmonization can resolve terminology issues where concepts and terms have developed differently in individual languages or language communities. This same can be said for emerging subject fields where concepts are developing and terms and definitions appearing in literature are, as yet, inconsistent. ISO 860:1996 [5] specifies a methodology for international harmonization of concepts, definitions, and terms.

The overall objective of applying the methodologies outlined in [4, 5] is to obtain a vocabulary in which a single term corresponds to a single concept, and conversely, a single concept corresponds to a single term. Moreover, definitions should be precise and non-circular, while terms should be concise and linguistically correct – prerequisites for improving the efficiency of communication in the subject field.

Harmonized Biometric Vocabulary

In 2002, the standards body ISO/IEC JTC 1 established Subcommittee 37 for the purpose of developing standards in the field of Biometrics. As is the case with many of the JTC 1 Subcommittees, a working group, Working Group (WG) 1 was established within Subcommittee (SC) 37 to develop a common vocabulary for use within the developing biometric standards projects. WG 1 Harmonized Biometric Vocabulary has the following Terms of Reference:

1. Concepts, terms, and definitions should be documented to be used throughout SC 37 International Standards, interacting as needed, with other SC 37 WGs.
2. Concepts, terms, and definitions should be articulated based on appropriate ISO/IEC standards for the development of ISO terminology.
3. Sources of terms and definitions should be identified for possible use in an SC 37 vocabulary (e.g., those drawn from existing standards, as well as other sources).
4. Ambiguity in terms and definitions in SC 37 Standards can be minimized which arise from the differences in cultures and languages.
5. The support and participation of experts to promote and progress the objectives and activities of SC 37/WG 1 should be identified and enlisted.
6. A standard on biometric vocabulary should be developed based on the concepts, terms, and definitions developed above to be proposed as a part of the ISO 2382 multi-part standard.
7. It serves as a source of expertise in this field to other WGs of SC 37. SC 37's goals can be supported by responding in a timely fashion to requests pertaining to the area of expertise in biometric vocabulary initiated by SC 37, its WGs or other organizations such as JTC 1 SCs, ISO TCs and other SC 37 Liaison organizations.

ISO 2382 [6] is a multi-part standard containing vocabulary developed in various ISO/IEC JTC 1 Subcommittees. SC 37 has reserved ISO 2382 Part 37 for publication of the completed biometrics vocabulary from WG 1.

Since its inception, the members of WG 1 have worked to harmonize and refine a biometrics vocabulary that follows the ISO guidelines for terminology work noted above. Biometrics is a relatively new subject field and thus biometrics literature tends to contain a variety of definitions for any single biometric term, as well as a variety of terms for seemingly the same concept. For example, consider the following:

Template/reference template [2]: Data, which represents the biometric measurement of an enrollee, used by a biometric system for comparison against subsequently submitted biometric samples.

Template [7]: A user's stored reference measure based on biometric feature(s) extracted from biometric sample(s).

Template [8]: A shortened term for a biometric *reference template.*

Reference template [8]: Also referred to as simply a *template*, the data in a biometric security system that represents the biometric measurement of a specific person's identity.

Template [3]: A mathematical representation of biometric data.

▶ *Note that **biometric data** is not defined in the FindBiometrics glossary, but **biometric** as an adjective is defined as: Of or pertaining to technologies that utilize behavioral or physiological characteristics to determine or verify identity.*

While each of these definitions appears to refer to the same concept, different characteristics are introduced into the definitions and they are:

- Data in a biometric security system
- Data used by a biometric system
- Mathematical representation
- Represents the biometric measurement of a specific person's identity
- Used for comparison against subsequently submitted biometric samples
- Stored

In addition, two different terms are presented *reference template* and *template* for the single concept. Many such examples permeate biometric literature:

Biometric feature [7]: A representation from a biometric sample extracted by the extraction system

Biometric data [2]: The information extracted from the biometric sample and used either to build a reference template (template data) or to compare against a previously created reference template (comparison data)

Feature extraction [3]: The automated process of locating and encoding distinctive characteristics from a biometric sample to generate a template

Although [FB] does not include the definition of *feature* in the glossary, one might infer the following from the above definition, *Feature*: distinctive characteristics from a biometric sample

It is easy to see how such diversity in designations and definitions for a single concept can compromise effective communication.

To resolve this, WG1 has collected terms and definitions from a variety of sources and continues to work on harmonizing the concepts and terms according to the guidelines in [5]. For example, WG1 has harmonized the above as described subsequently.

Biometric Template

Characteristics

- Stored Biometric features
- Attributed to an individual at enrollment
- Type of biometric reference
- Comparison uses a function not dependent on individual e.g., Hamming distance, Euclidean distance, etc., although it's parameters might be
- Directly compared to sample features
- Set of features that can be compared directly to the input features to give a score

Definition

Set of stored biometric features comparable directly to biometric features of a recognition of biometric sample

NOTE 1 A biometric reference consisting of an image, or other captured biometric sample, in its original, enhanced or compressed form, is not a biometric template.

NOTE 2 The biometric features are not considered to be a biometric template unless they are stored for reference.

Biometric Feature

Characteristics

- Output of a completed biometric feature extraction process
- Numbers or labels extracted from biometric samples and used for comparison

Definition

Numbers or labels extracted from biometric samples and used for comparison

NOTE 1 Biometric features are the output of a completed biometric feature extraction process.

NOTE 2 The use of this term should be consistent with its use by the pattern recognition and mathematics communities.

NOTE 3 A biometric feature set can also be considered as a processed biometric sample.

The biometric vocabulary under development within WG 1 is continually updated in Standing Document 2 [9]. Since the development process is an iterative process, existing concepts will continue to be refined as relationships among concepts are explored and new concepts are introduced. ISO/IEC JTC1 SC 37

Standing Document 2 [9] is broken into two parts, the main body which includes terms and definitions for concepts that have been harmonized, and a series of concept diagrams included as annexes to demonstrate the relations among concepts and to illustrate concepts that are still to be developed.

Given that ISO/IEC JTC 1 SC37 is an international organization, the members of WG 1 represent several countries, including, Canada, France, Germany, Japan, Singapore, Spain, South Africa, the Russian Federation, and the United Kingdom. As a result, the translatability of terms and definitions into various languages is considered throughout as the harmonization process. The Russian Federation National Body has provided a first draft Russian translation of the terms and definitions of Standing Document 2. A German translation has also been developed [10], and will be updated as the Standing Document 2 evolves. As Standing Document 2 eventually becomes published to ISO 2382 Part 37, it is the hope of WG1 to include at least Russian, French, and German translations.

Related Entries

- ▶ Biometric Data Interchange Format, Standardization
- ▶ Biometric Sample Quality, Standardization
- ▶ Biometric Technical Interface, Standardization
- ▶ Biometrics Security, Standardization
- ▶ Performance Testing Methodology Standardization

References

1. ISO 1087-1:2000, *Terminology work — Vocabulary — Part 1: Theory and application.*
2. 1999 Glossary of Biometric Terms, International Association for Biometrics (iAfB): http://www.afb.org.uk/docs/glossary.htm
3. Find BIOMETRICS Glossary: http://www.findbiometrics.com/Pages/glossary.html
4. ISO 704:2000, *Terminology work — Principles and methods.*
5. ISO 860:1996, *Terminology work — Harmonization of concepts and terms.*
6. ISO 2382, *Information Technology – Vocabulary.*
7. Common Criteria Biometric Evaluation Methodology, v1.0: http://www.cesg.gov.uk/site/ast/biometrics/media/BEM_10.pdf
8. Wikipedia: http://wikipedia.org/
9. ISO/IEC JTC1 SC 37 Standing Document 2 – *Harmonized Biometric Vocabulary*: http://isotc.iso.org/livelink/livelink?func=ll&objId=2299739&objAction=browse&sort=name
10. German translation of SD2: http://www.3dface.org/media/vocabulary.html

Biometric Vulnerabilities, Overview

Andy Adler[1], Stephanie Schuckers[2]
[1]Carleton University, Ottawa, ON, Canada
[2]Clarkson University, Potsdam, NY, USA

Definition

Biometric systems, like all security systems, have vulnerabilities. This article provides a survey of the many possibilities of attack against traditional biometric systems. The vulnerabilities of nontraditional systems, such as those based on encoded biometrics are surveyed in the chapter *Security and Liveness: Overview*. Here, biometric system security is defined by its absence: a vulnerability in biometric security results in incorrect recognition or failure to correctly recognize individuals. This definition includes methods to falsely accept an individual (spoofing) impact overall system performance (denial of service), or to attack another system via leaked data (identity theft). In this chapter, each stage of biometrics processing is analyzed and the potential vulnerabilities are discussed. Techniques to structure the analysis of vulnerabilities, such *Attack Trees* are described, and four application scenarios and their vulnerabilites are considered.

Introduction

This chapter surveys the many types of security vulnerabilities in traditional biometric systems. For a more general survey of security issues in biometric systems, including those for novel and encrypted biometric schemes, ▶ Security and Liveness, Overview. Biometric system vulnerabilities are defined as avenues of attack against a biometric system that involve an active attacker. The resistance of a biometric system to zero-effort attack is the system false accept rate (FAR), and this value is generally considered to be the *performance* of the biometric system. Since there are many configurations for biometric systems and many possible ways to attack each, the topic of biometric system vulnerabilities is necessarily very broad; this chapter describes classes of biometric applications and reviews the vulnerabilities of each.

Note that this chapter concentrates on system vulnerabilities, which are part of the biometric processing

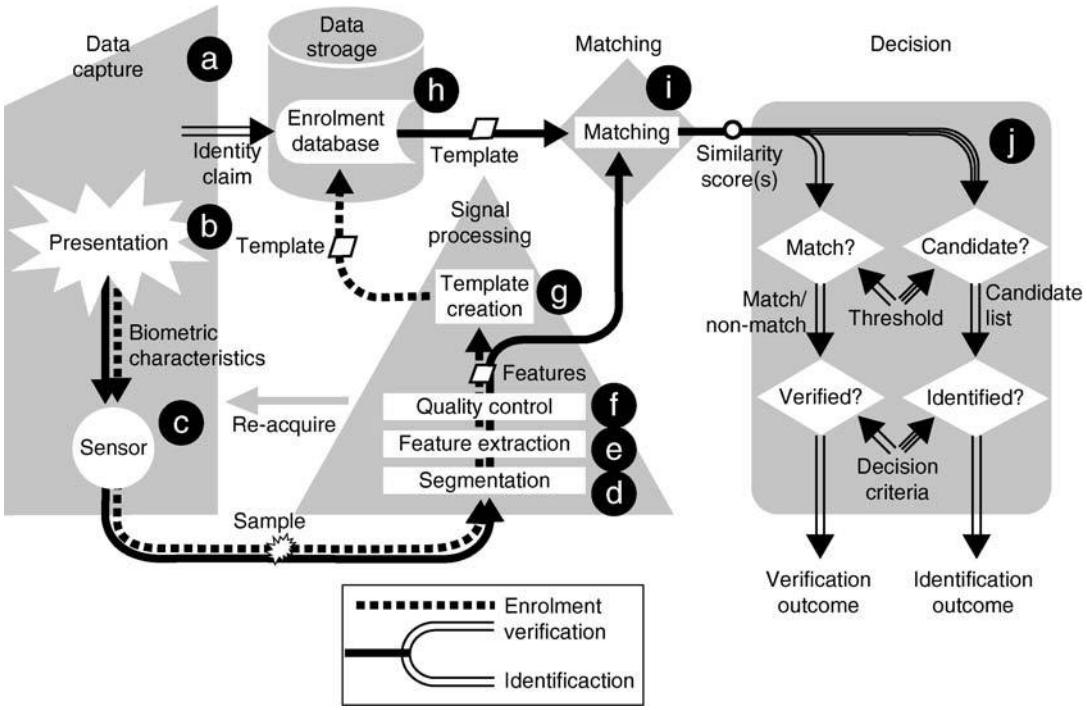
itself. Since biometric systems are implemented on server computers, they are vulnerable to all cryptographic, virus, and other attacks, which plague modern computer systems [1]; These issues have been pointed out, but not covered in detail.

Biometric Subsystems and Vulnerabilities

To classify biometric security vulnerabilities, it is typical to study each subsystem and interconnection in a system diagram (Fig. 1). Early work is presented in [3], with later contributions coming from [4, 5, 6]. Each system module is considered in turn.

Identity Claim (A)

Identity claims are not biometric properties, but form an essential part of most biometric security systems. Exceptions are possible: an example is verifying a season ticket holder; the person's identity doesn't matter, as long as they have paid. Identity claims are primarily based on links to government issued identity documents, and are thus vulnerable to all forms of fraud of such documents. This is a problem even for highly secure documents, such as passports, which are often issued on the basis of less secure "breeder documents" [7], such as birth certificates issued by local government, hospital, or even religious authorities.



Biometric Vulnerabilities, Overview. Figure 1 Biometric System Block Diagram (from [2]). Steps **a** – **h** are analyzed in detail in this chapter. Each presented sample (**b**) is acquired by a sensor (**c**) processed via segmentation (**d**) and feature extraction (**d**) algorithms. If available, a sample quality (**e**) assessment algorithm is used to indicate a need to reacquire the sample. Biometric features are encoded into a template, which is stored (**h**) in a database, on an identity card or in secure hardware. For biometric encryption systems, a code or token is combined with the biometric features in the template. During enrollment, biometric samples are linked to a claimed identity (**a**), and during subsequent verification or identification, samples are tested against enrolled samples, using a matching algorithm (**i**) and an identity decision (**J**) is made, either automatically, or by a human agent reviewing biometric system outputs.

Presentation (B)

An attack on the biometric sensor provides false biometric sample into the system. Such attacks are designed to either avoid detection (false negative) or masquerade as another (false positive). The latter attack is typically called spoofing. Clearly, avoiding detection is easier than masquerading, since features simply need to be changed enough to confuse the segmentation or feature extraction module. Changing makeup, facial hair, and glasses or abrading or wetting fingers is often successful; although recent progress in biometric algorithms has reduced the effectiveness of such techniques. Knowledge of the details of algorithms can make such attacks easier; for example, rotating the head will confuse many iris algorithms that do not expect image rotation of more than a few degrees.

An attempt to gain unauthorized access using presentation of an artificial biometric, which copies that of an authorized user is called a "spoof". The most well known spoofs are for fingerprint; it is possible to spoof a variety of fingerprint technologies through relatively simple techniques using casts of a finger with molds made of household materials [8, 9]. A morbid concern is the use of dismembered fingers, which can be scanned and verified against enrolled fingers. Other modalities may be spoofed: face using pictures or high resolution video, iris with contact lenses, and voice recordings for voice biometrics [9]. Techniques to make spoofing more difficult include *liveness*, multiple biometrics, and use of biometrics in combination with a challenge response, passwords, tokens, or smart cards. The goal of liveness testing is to determine if the biometric being captured is an actual measurement from a live person who is present at the time of capture [10]. Typically, liveness is a secondary measure after biometric authentication, which must be needed to achieve a positive response. Liveness may be implemented in hardware or software. Hardware liveness tests require additional sensors in conjunction with the biometric sensor, increasing cost. Examples of this approach include thermal sensing of finger temperature, ECG, impedance of the skin, and pulse oximetry. Software liveness tests rely of further processing of the biometric signal to gather liveness information [11]. Examples include quantifying saccade movements in the eye for iris recognition, lip-reading, or perspiration in the fingerprint [12]. It is important to note that the liveness measurement increases the difficulty of – but does not prevent – fraudulent presentation. Furthermore, liveness stage may have vulnerabilities, for example, using a translucent spoof in combination with a live finger to fool pulse oximetry.

Sensor (C)

Attacks on the biometric sensor include any technique that subverts or replaces the sensor hardware. In some cases subverting the sensor allows complete bypassing of the biometric system. For example, in some biometric door locks, the sensor module includes the entire biometric system including a Wiegand output or relay output to activate the solenoid in a door lock. Subverting such a system may be as simple as physically bypassing the biometric system.

In many cases, an attack on the sensor would take the form of a replay. The connection between the biometric sensor and the biometric system is subverted to allow input of arbitrary signals, and images from legitimate users are input into the system. To obtain the signals, several strategies may be employed. Eavesdropping requires hiding the recording instruments and wiring of the sensor. For biometrics using contactless smart cards, such eavesdropping becomes more feasible (see [13]). Another approach is to record signals from a sensor under the control of the attacker.

Protection of the sensor typically requires cryptographic techniques to prevent capture and relay of signals and replacement of the sensor [1]. This imposes a larger cost for sensors with integrated cryptographic capability and for management of the security and key infrastructure.

Segmentation (D)

Biometric segmentation extracts the image or signal of interest from the background, and a failure to segment means the system does not detect the presence of the appropriate biometric feature. Segmentation attacks may be used to escape surveillance or to generate a denial of service (DoS) attack. For example, consider a surveillance system in which the face detection algorithm assumes faces have two eyes. By covering an eye, a person is not detected in the biometric system.

Another example would be where parts of a fingerprint core are damaged to cause a particular algorithm to mislocate the core. Since the damaged area is small, it would not arouse the suspicion of an agent reviewing the images.

Feature Extraction (E)

Attacks of the feature extraction module can be used either to escape detection or to create impostors. The first category is similar to those of Segmentation. Knowledge of the feature extraction algorithms can be used to design special features in presented biometric samples to cause incorrect features to be calculated.

Characterizing feature extraction algorithms: To implement such an attack, it is necessary to discover the characteristics of the feature extraction algorithm. Are facial hair or glasses excluded (face recognition)? How are the eyelid/eyelash regions detected and cropped (iris recognition)? Most current high performing biometric recognition algorithms are proprietary, but are often based on published scientific literature, which may provide such information. Another approach is to obtain copies of the biometric software and conduct offline experiments. Biometric algorithms are likely susceptible to reverse engineering techniques. It would appear possible to automatically conduct such reverse engineering, but we are not aware of any published results.

Biometric "zoo": There is great variability between individuals in terms of the accuracy and reliability of their calculated biometric features. Doddington et al. developed a taxonomy for different user classes [14]. *Sheep* are the dominant type, and biometric systems perform well for them. *Goats* are difficult to recognize. They adversely affect system performance, accounting for a significant fraction of the FRR. *Lambs* are easy to imitate – a randomly chosen individual is likely to be identified as a lamb. They account for a significant fraction of the FAR. *Wolves* are more likely to be identified as other individuals, and account for a large fraction of the FAR. The existence of lambs and wolves represents a vulnerability to biometric systems. If wolves can be identified, they may be recruited to defeat systems; similarly, if lambs can be identified in the legitimate user population, either through correlation or via directly observable characteristics, they may be targets of attacks.

Quality Control (F)

Evaluation of biometric sample quality is important to ensure low biometric error rates. Most systems, especially during enrollment, verify the quality of input images. Biometric quality assessment is an active area of research, and current approaches are almost exclusively algorithm specific. If the details of the quality assessment module can be measured (either though trial and error or through off-line analysis), it may be possible to create specific image features, which force classification in either category. Quality assessment algorithms often look for high frequency noise content in images as evidence of poor quality, while line structures in images indicate higher quality. Attacks on the quality control algorithm are of two types: classifying a good image as poor, and classifying a low quality image as good. In the former case, the goal of the attack would be to evade detection, since poor images will not be used for matching. In the latter case, low quality images will be enrolled. Such images may force internal match thresholds to be lowered (either for that image, or in some cases, globally). Such a scenario will create "lambs" in the database and increase system FAR.

Template Creation (G)

Biometric features are encoded into a template, a (proprietary or standards-conforming) compact digital representation of the essential features of the sample image. One common claim is that, since template creation is a one-way function, it is impossible or infeasible to regenerate the image from the templates [15]; however, it has been shown that it is generally possible to regenerate versions of biometric sample images from templates [16]. These regenerated images may be used to masquerade at the sensor or to generate a spoofed biometric for presentation (▶ Template security).

Interoperability: Government applications of biometrics need to be concerned with interoperability. Biometric samples enrolled on one system must be usable on other vendor systems if a government is to allow cross-jurisdictional use, and to avoid vendor lock-in. However, recent work on interoperability has revealed it to be difficult, even when all vendors conform to standards. Tests of the International Labor Organization seafarer's ID card [17] showed

incompatibilities with the use of the minutiae type "other" and incompatible ways to quantize minutiae angles. Such interoperability difficulties present biometric system vulnerabilities, which could be used to increase FRR or for a DoS attack.

Data Storage (H)

Enrolled biometric templates are stored for future verification or identification. Vulnerabilities of template storage concern modifying the storage (adding, modifying or removing templates), copying template data for secondary uses (identity theft or directly inputting the template information at another stage of the system to achieve authentication), or modifying the identity to which the biometric is assigned.

Storage may take many forms, including databases (local or distributed), on ID documents (into a smart card [13], or 2D barcode [17]), or on electronic devices (a hardened token [18], laptop, mobile telephone, or door access module). Template data may be in plaintext, encrypted or digitally signed. In many government applications, it may be necessary to provide public information on the template format and encryption used, to reassure citizens about the nature of the data stored on their ID cards, but this may also increase the possibility of identity theft. Vulnerabilities of template storage are primarily those of the underlying computer infrastructure, and are not dealt with in detail here.

Template transmission: The transmission medium between the template storage and matcher is similarly vulnerable to the template storage. In many cases, attacks against template data transmission may be easier than against the template storage. This is especially the case for passive eavesdropping and recording of data in transit for wireless transmission (such as contactless ID cards). Encrypted transmission is essential, but may still be vulnerable to key discovery [13].

Matching (I)

A biometric matcher calculates a similarity score related to the likelihood that two biometric samples are from the same individual. Attacks against the matcher are somewhat obscure, but may be possible in certain cases. For biometric fusion systems, extreme scores in one biometric modality may override the inputs from other modalities. Biometric matchers, which are based on Fisher discriminant strategies calculate global thresholds based on the between class covariance, which may be modified by enrolling specifically crafted biometric samples.

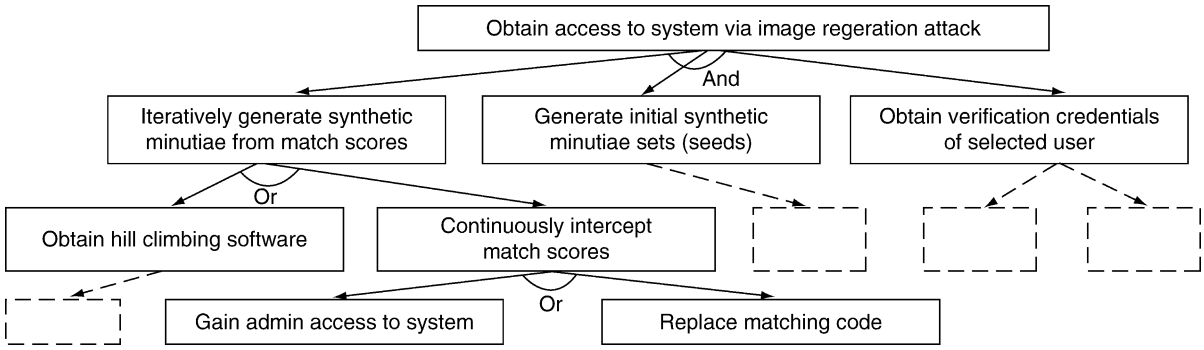
Decision (J)

Biometric decisions are often reviewed by a human operator (such as for most government applications). Such operators are well known to be susceptible to fatigue and boredom. One of the goals of DoS attacks can be to force operators to abandon a biometric system, or to mistrust its output (by causing it to produce a sufficiently large number of errors) [1].

Attack Trees

Complex systems are exposed to multiple possible vulnerabilities, and the ability to exploit a given vulnerability is dependent on a chain of requirements. Vulnerabilities vary in severity, and may be protected against by various countermeasures, such as: supervision of enrollment or verification, liveness detection, template anonymization, cryptographic storage and transport, and traditional network security measures. Countermeasures vary in maturity, cost, and cost-effectiveness. To analyze such a complex scenario, the factors may be organized into *attack trees*. This analysis methodology was developed by Schneier [19] and formalized by Moore et al. [20]. In [19], the example attack "Open Safe," is analyzed to occur due to "Pick Lock", "Learn Combo," "Cut Open Safe," or "Install Improperly." "Learn Combo" may, in turn, occur due to "Eavesdrop", "Bribe," or other reasons, which in turn depend on further factors. The requirements for each factor can be assessed (Eavesdropping requires a technical skill, while bribing requires an amount of money). Attack trees may be analyzed by assigning each node with a feasibility, the requirement for special equipment, or cost.

Attack tree techniques for biometric system security have been developed by Cukic and Barlow [4]. Figure 2 shows a fraction of the attack tree [4] for image regeneration from templates [6].



Biometric Vulnerabilities, Overview. Figure 2 Attack tree fraction adapted from [4] (dotted blocks represent removed tree portions) to implement the template regeneration attack of [6]. AND/OR nodes indicate that *all*/*one* of the sub-blocks are/is required.

Application Profiles and Vulnerabilities

This chapter reviews a large list of possible vulnerabilities in biometric security systems. Such a large list can perhaps give the impression that biometric systems are extremely insecure. In this context, it is important to clarify that many potential vulnerabilities are not a concern in many biometric applications. For example, in a particular application, if security is one of the primary reasons for choosing a biometric (over, say, convenience), it is also important to look at the context of the security mechanism it is replacing. One could certainly argue that biometrically enabled passwords (even with weaknesses discussed as below) have improved security in this application over conventional passports.

To clarify the security requirements of various biometric implementations, four different biometric application scenarios are considered: government ID cards, physical access, computer and network access, and digital content protection.

Government Identity Cards

Perhaps the most widely discussed applications for biometrics are for government identity cards. For example, the new ICAO machine readable passport standards require biometric data in passports. Passports have an embedded contactless smart card, into which face recognition (mandatory) and fingerprint or iris (optional) biometric templates are stored encrypted in a standardized format.

To allow data interchange, the encryption key is based on information available in the machine readable zone. A recent report has demonstrated the ability to contactlessly read the new UK passports [13]. This raises the concern that biometric and biographical data may be surrepitiously copied and used for identity theft. Biometric enabled passports have been strongly criticized by privacy advocates (e.g., [21]). Given the privacy concerns associated with a large government database, several authors have questioned whether the additional security is worth it [7].

Government ID applications of biometrics are characterized by the following requirements and concerns:

- *Interoperability and standards compliance.* Interoperability is difficult to achieve for complex systems such as biometrics (e.g., [17]); poorly interoperable systems give poor performance and are vulnerable to attacks such as denial of service.
- *Cryptographic compability.* To allow interchange of encrypted documents, public key cryptographic systems are required, in which the public keys are made available to receiving governments. Considering the wide distribution of keys, it must be assumed that the public keys will be fairly easily available to attackers.
- *Large databases of vulnerable data.* Identity document data is typically stored in large centralized databases; however, these become vulnerable, and high value targets for attack. Several high profile cases of compromise of government databases have occurred.
- *Secondary use of government IDs.* Government identity cards often have secondary uses; for

example, driver's licenses are used to prove name, age, and even citizenship. This means that biometric documents designed for a narrow range of security concerns may be used in very different threat environments, with inadvertent side effects.
- *Typically supervised use.* For most applications of government biometric identity, the point of application will be supervised (e.g., immigration control). This makes spoofing more difficult for these applications.

Physical Access

Physical access systems for biometrics are typically for government and industrial applications. In "time and attendance systems" biometrics measure arrival and departure times of staff. In physical access security systems, secure spaces are controlled by biometric sensors. These spaces may be an entire site, or restricted parts of a worksite.

Physical access applications are characterized by the following requirements and concerns:

- *Concern about privacy.* Staffs are often concerned that biometric records will be controlled by the employer and may be provided to police. It is important to address this concern both technically, and by clear communication with staff.
- *Unsupervised sensors.* Physical access sensors are typically unsupervised. This means that there is a potential vulnerability to spoofing and other attacks at the presentation and sensor.
- *Workarounds.* It is well known that busy staff see security as a burden to work around. Biometrics has the advantage that staff often see it as more convenient than keys or identity cards, encouraging compliance. However, if the system is implemented in a cumbersome way, there is an incentive to work around burdensome infrastructure, by propping open doors, etc.

Computer and Network Access

Biometric system can facilitate secure access to computer systems and networks; this is an important requirement in government, health care, and banking applications, as well as many others. Biometric sensors have recently been provided with many laptop computer systems. These applications, characterized by the following requirements and concerns:

- *Assurance levels.* The biometric system security needs to be matched to the security level (or assurance level) of the overall system. An excellent review of the security of biometric authentication systems is [18]. Each assurance level from "passwords and PINs" to "Hard crypto token" is analyzed to determine whether (and which type of) biometric devices are suitable.
- *Network attacks.* Biometric systems for network access are vulnerable to many of the attacks, which can be mounted across a computer network. Examples are relay of issued credentials, and virus and other security compromises of the desktop computers (to which biometrics are often attached). Security must, therefore, include computer security and cryptographic protection of biometric data and security tokens.
- *Password caching.* Most biometric software solutions do not actually replace passwords, but simply keep a cache of security keys. A valid biometric sample will make the software search for the appropriate key to unlock the application. However, this means that cracking the software will release both the security keys, and the biometric template of the user.

Digital Content Protection

Biometrics have been considered as a way to protect copyright content, such as music and videos. In such a scenario, the content is encrypted and bound to the biometric of the purchaser [22]. It may be assumed that biometrically locked digital documents will be subject to attacks, especially since both the documents and the software to access them will be widely distributed [22]. These applications, characterized by the following concerns:

- *Incentive to crack systems.* Digital content protection systems are under the control of an (often hostile) user population, which creates and incentive to crack the security systems. Additionally, any such security breaches tend to be published on the internet resulting in wide scale use and potential poor publicity for the content providers.

- *Privacy and identity theft concerns.* Locking of digital content with biometrics tends to create concerns about privacy among users, since breaches of the security can potentially compromise the biometric security for large numbers of users.

Summary

This chapter provides a broad overview of vulnerabilities in biometric systems. Vulnerabilities are defined in terms of possible active attacks against biometric systems. A model of biometric processing [2] is considered in detail, and the potential vulnerabilities at each stage of processing are considered: identity claim, presentation, sensor, segmentation, feature extraction, quality control, template creation, data storage, matching, and decision. To understand the vulnerabilities of a large biometric system, attack tree methods are explained. Finally, four example scenarios are given for biometric applications, the vulnerabilities are considered: government identity cards, physical access, computer and network access, and digital content protection. However, in addition to the vulnerabilities specific to the biometric technology, it is important to note that the vulnerabilities of any networked computer security system continue to be a concern; specifically, such systems are vulnerable to *social engineering* and all the security issues which plague modern computer networks. Finally, biometric vulnerabilities must be compared to those of the systems they are designed to replace. In many cases, the biometric system, with the vulnerabilities considered in this chapter, will still be dramatically more secure than identity cards, passwords, or other tokens.

Related Entries

- ▶ Biometric Encryption
- ▶ Biometric Security, Overview
- ▶ Biometric System Design, Overview
- ▶ Biometrics Security, Standardization
- ▶ Cancelable Biometrics
- ▶ Fraud Reduction, Application
- ▶ Fraud Reduction, Overview
- ▶ Security Issues, System Design
- ▶ Tampler-Proof OS Zero-Effort Forgery Test

References

1. Ferguson, N., Schneier, B.: Practical Cryptography. Wiley, NJ, USA (2003)
2. ISO: Standing Document 2, version 5 – Harmonized Biometric Vocabulary. Technical Report ISO/IEC JTC 1/SC 37 N 1480 (2006)
3. Ratha, N.K., Connell, J.H., Bolle, R.M.: Enhancing security and privacy in biometrics-based authentication systems. IBM Syst. J. **40**, 614–634 (2001)
4. Cukic, B., Barlow, N.: Threats and countermeasures, In Proc. Biometric Consortium Conference, Washington DC, USA (2005)
5. Tilton, C: Biometrics in E-Authentication: Threat model. Biometrics Consortium Conference, Baltimore, MD, USA (2006)
6. Uludag, U., Jain, A.K.: Attacks on biometric systems: A case study in fingerprints. In Proc. SPIE-EI 2004, Security, Steganography and Watermarking of Multimedia Contents VI **5306,** 622–633 (2004)
7. Salter, M.B.: Passports, mobility, and security: How smart can the border be?. Int. Stud. Persp. **5**, 71–91 (2004)
8. Matsumoto, T., Matsumoto, H., Yamada, K., Hoshino, S.: Impact of artificial "gummy" fingers on fingerprint Systems. Proc SPIE, **4677,** January (2002)
9. Thalheim, L., Krissler, J.: Body check: Biometric access protection devices and their programs put to the test. ct magazine, November (2002)
10. International Biometric Group, Liveness Detection in Biometric Systems, http://www.ibgweb.com/reports/public/reports/liveness.html
11. Daugman, J.: Iris recognition and spoofing countermeasures. 7th Int. Biometric Conference, London (2002)
12. Derakhshani, R., Schuckers, S.A.C., Hornak, L.A., O'Gorman, L.: Determination of vitality from a non-invasive biomedical measurement for use in fingerprint scanners. Pattern Recognition **36**, 386–396 (2003)
13. The Guardian (17 Nov. 2006) Cracked it!
14. Doddington, G., Liggett, W., Martin, A., Przybocki, N., Reynolds, D.: Sheep, goats, lambs and wolves: an analysis of individual differences in speaker recognition performance. In Proc. Int. Conf. Auditory-Visual Speech Processing, Sidney, Australia (1998)
15. International Biometric Group: Generating images from templates. http://www.ibgweb.com/reports/public/reports/templates_images.html (2002)
16. Jain, A.K., Nagar, A., Nandakumar, K.: Biometric template security. EURASIP. J. Adv. Signal. Proc. chapter ID 579416, 17 (2008)
17. International Labour Organization: Biometric Testing Campaign Report (Addendum to Part I). Geneva (2005)
18. International Committee for Information Technology Standards (INCITS): Study Report on Biometrics in E-Authentication, Technical Report INCITS M1/06-0693 (2006)
19. Schneier, B.: Attack trees. Dr. Dobb's J. (1999)
20. Moore, A.P., Ellison, R.J., Linger, R.C.: Attack Modeling for Information Security and Survivability. Carnegie Mellon University, Pittsburgh, PA, USA (2001)

21. Ross, P.E.: Loser: passport to nowhere. IEEE Spectrum **42**, 54–55 (2005)
22. Kundur, D., Lin, C.-Y., Macq, B., Yu, H.: Special Issue on enabling security technologies for digital rights management. Proc. IEEE **92**, 879–882 (2004)

Biometric Watermarking

► Iris Digital Watermarking

Biometrics, Overview

Arun Ross[1], Anil K. Jain[2]
[1]Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV, USA
[2]Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA

Synonyms

Biometric system; Biometric recognition

Definition

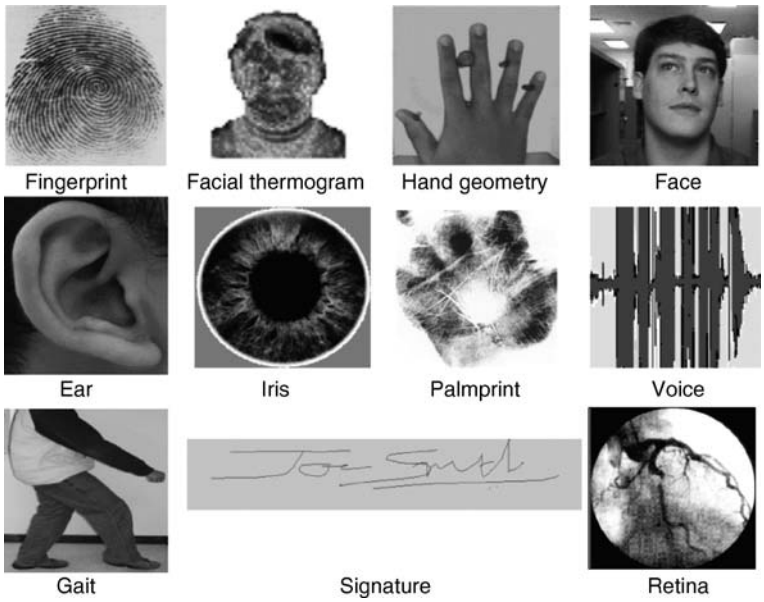
Biometrics is the science of establishing the identity of a person based on the physical (e.g., fingerprints, face, hand geometry, and iris) or behavioral (e.g., gait, signature, and keyboard dynamics) attributes associated with an individual. A typical biometric system uses appropriately designed sensors to capture the biometric trait of a person and compares this against the information stored in a database to establish identity. A biometric system can operate in two distinct modes: in the verification mode, the system *confirms or negates* a claimed identity, while in the identification mode, it *determines* the identity of an individual.

Introduction

A wide variety of systems require reliable authentication schemes to confirm the identity of an individual requesting their services. The purpose of such schemes is to ensure that the rendered services are accessed only by a legitimate user, and not anyone else. Examples of such applications include secure access to buildings, computer systems, laptops, cellular phones, and ATMs. In the absence of robust authentication schemes, these systems are vulnerable to the wiles of an impostor.

Traditionally, passwords (knowledge-based security) and ID cards (token-based security) have been used to restrict access to systems. However, security can be easily breached in these systems when a password is divulged to an unauthorized user or an ID card is stolen by an impostor. Further, simple passwords are easy to guess (by an impostor) and complex passwords may be hard to recall (by a legitimate user). The emergence of *biometrics* has addressed the problems that plague these traditional security methods. Biometrics refers to the automatic identification (or verification) of an individual (or a claimed identity) by using certain physical or behavioral traits associated with the person. By using biometrics, it is possible to establish an identity based on "who you are," rather than by "what you possess" (e.g., an ID card) or "what you remember" (e.g., a password). Current biometric systems make use of fingerprints, hand geometry, iris, retina, face, hand vein, facial thermograms, signature, voiceprint, etc. (Fig. 1) to establish a person's identity [1–5]. While biometric systems have their limitations (e.g., additional cost, temporal changes in biometric traits, etc.), they have an edge over traditional security methods in that they cannot be easily stolen, shared, or lost.

Biometric systems also introduce an aspect of user convenience that may not be possible using traditional security techniques. For example, users maintaining different passwords for different applications may find it challenging to recollect the password associated with a specific application. In some instances, the user might even forget the password, requiring the system administrator to intervene and reset the password for that user. Maintaining, recollecting, and resetting passwords can, therefore, be a tedious and expensive task. Biometrics, however, addresses this problem effectively: a user can use the same biometric trait (e.g., right index finger) or different biometric traits (e.g., fingerprint, hand geometry, iris) for different applications, with "password" recollection not being an issue at all.



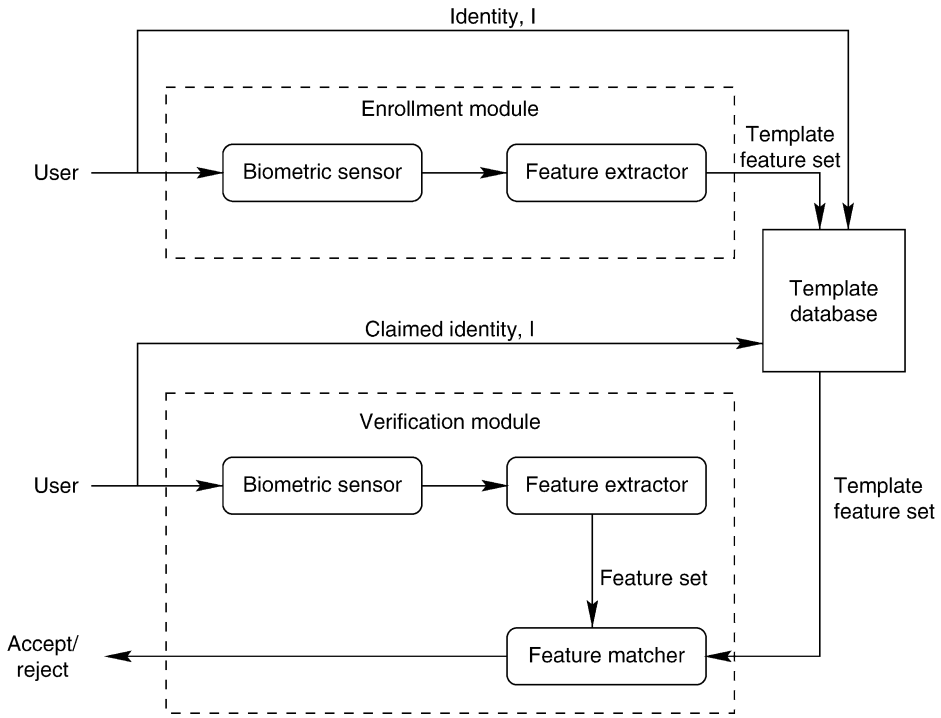
Biometrics, Overview. Figure 1 Examples of some of the biometric traits used for authenticating an individual.

Operation of a Biometric System

A typical biometric system operates by acquiring biometric data from an individual, extracting a feature set from the acquired data, and comparing this feature set against the ▶ template feature set stored in the database (Fig. 2). In an *identification* scheme, where the goal is to recognize the individual, this comparison is done against templates corresponding to all the enrolled users (a one-to-many matching); in a *verification* scheme, where the goal is to verify a claimed identity, the comparison is done against only those templates corresponding to the claimed identity (a one-to-one matching). Thus, identification ("Whose biometric data is this?") and verification ("Does this biometric data belong to Bob?") are two different problems with different inherent complexities. The templates are typically created at the time of enrollment, and depending on the application, may or may not require human personnel intervention.

Biometric systems are being increasingly deployed in large scale civilian applications. The Schiphol Privium scheme at the Amsterdam airport, for example, employs iris scan cards to speed up the passport and visa control procedures. Passengers enrolled in this scheme insert their card at the gate and look into a camera; the camera acquires the image of the traveler's eye and processes it to locate the iris, and compute the Iriscode; the computed Iriscode is compared with the data residing in the card to complete user verification. A similar scheme is also being used to verify the identity of Schiphol airport employees working in high-security areas. Thus, biometric systems can be used to enhance user convenience while improving security.

A simple biometric system has four important modules [6]: (1) *Sensor module* which acquires the biometric data of an individual. An example would be a fingerprint sensor that images the fingerprint ridges of an user; (2) *Feature extraction module* in which the acquired biometric data is processed to extract a feature set that represents the data. For example, the position and orientation of ridge bifurcations and ridge endings (known as minutiae points) in a fingerprint image are extracted in the feature extraction module of a fingerprint system; (3) *Matching module* in which the extracted feature set is compared against that of the template by generating a match score. For example, in this module, the number of matching minutiae points between the acquired and template fingerprint images is determined, and a matching score reported. (4) *Decision-making module* in which the user's claimed identity is either accepted or rejected based on the matching score (verification). Alternatively, the system may identify an user based on the matching scores (identification).



Biometrics, Overview. **Figure 2** The enrollment module and the verification module of a biometric system.

Quantifying Performance

Unlike password-based systems, where a perfect match between two alphanumeric strings is necessary to validate a user's identity, a biometric system seldom encounters two samples of a user's biometric trait that result in exactly the same feature set. This is due to imperfect sensing conditions (e.g., noisy fingerprint due to sensor malfunction), alterations in the user's biometric characteristic (e.g., respiratory ailments impacting speaker recognition), changes in ambient conditions (e.g., inconsistent ▶ illumination levels in face recognition), and variations in the user's interaction with the sensor (e.g., occluded iris or partial fingerprints). The variability observed in the biometric feature set of an individual is referred to as *intra*class variation, and the variability between feature sets originating from two different individuals is known as *inter*class variation. A useful feature set exhibits small intra-class variation and large interclass variation.

A similarity match score is known as a genuine score or authentic score if it is the result of matching two samples of the same biometric trait of a user. It is known as an impostor score if it involves comparing two biometric samples originating from different users. To analyze the performance of a biometric system, the probability distribution of genuine and impostor matching scores is examined. A genuine matching score is obtained when two feature sets corresponding to the *same* individual are compared, and an impostor matching score is obtained when feature sets from two *different* individuals are compared. In the case of verification, when a matching score exceeds a certain threshold, the two feature sets are declared to be from the same individual; otherwise, they are assumed to be from different individuals. Thus, there are two fundamental types of errors associated with a verification system: (i) a false match, which occurs when an impostor matching score exceeds the threshold, and (ii) a false nonmatch, which occurs when a genuine matching score does not exceed the threshold. The error rates of systems based on fingerprint and iris are usually lower when compared to those based on voice, face, and hand geometry. A Receiver Operating Characteristic (ROC) curve plots the False Non-match Rate (FNMR – the percentage of genuine scores that do not exceed the threshold) against the False Match Rate (FMR – the percentage of impostor scores that exceed the threshold) at various thresholds. The operating threshold employed by a system depends on the nature of the application. In forensic applications, for example, a low FNMR is preferred, while in

high security access facilities like nuclear labs, a low FMR is desired (Fig. 3).

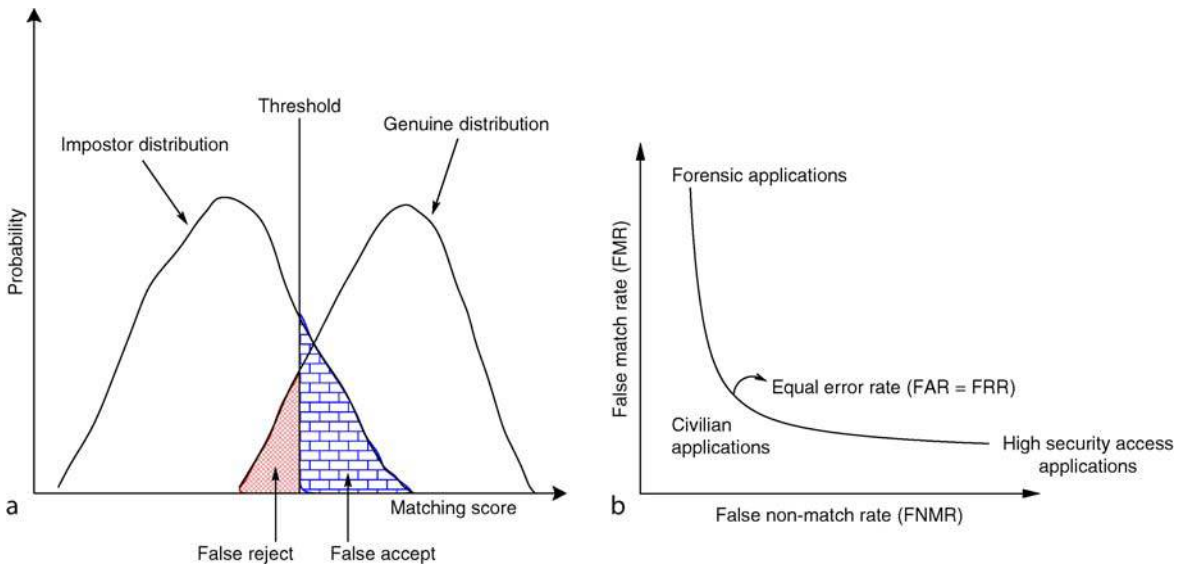
In the case of identification, the input feature set is compared against all templates residing in the database to determine the top match (i.e, the best match). The top match can be determined by examining the match scores pertaining to all the comparisons and reporting the identity of the template corresponding to the largest similarity score. The *identification rate* indicates the proportion of times a previously enrolled individual is successfully mapped to the correct identity in the system. Here, assume that the question being asked is, "Does the top match correspond to the correct identity?" An alternate question could be, "Does any one of the top k matches correspond to the correct identity?" (see [7]). The rank-k identification rate, R_k, indicates the proportion of times the correct identity occurs in the top k matches as determined by the match score. Rank-k performance can be summarized using the Cumulative Match Characteristic (CMC) that plots R_k against k, for $k = 1, 2, \ldots M$ with M being the number of enrolled users.

Besides FMR and FNMR, other types of errors are also possible in a biometric system. The Failure to Enroll (FTE) error refers to the inability of a biometric system to enroll an individual whose biometric trait may not be of good ▶ quality (e.g., poor quality fingerprint ridges). Similarly, a biometric system may be unable to procure good quality biometric data from an individual during authentication resulting in a Failure to Acquire (FTA) error.

A biometric system is susceptible to various types of attacks [8]. For example, an impostor may attempt to present a fake finger or a face mask or even a recorded voice sample to circumvent the system. The problem of fake biometrics may be mitigated by employing ▶ challenge-response mechanisms or conducting liveness detection tests. Privacy concerns related to the use of biometrics and protection of biometric templates are the issues that are currently being studied [9–11].

Applications

Establishing the identity of a person with high confidence is becoming critical in a number of applications in our vastly interconnected society. Questions like "Is she really who she claims to be?", "Is this person authorized to use this facility?" or "Is he in the watchlist posted by the government?" are routinely being posed in a variety of scenarios ranging from issuing a driver's license to gaining entry into a country. The need for reliable user



Biometrics, Overview. **Figure 3** Evaluating the matching accuracy of a biometric system operating in the verification mode. (**a**) Histograms of genuine and impostor matching scores and the two types of errors (False Accept and False Reject) that are possible in a verification system. (**b**) A Receiver Operating Characteristic (ROC) curve indicating the operating point (threshold) for different types of applications. Note that FMR and FNMR are often used as synonyms for FAR and FRR, respectively.

Biometrics, Overview. Table 1 Authentication solutions employing biometrics can be used in a variety of applications which depend on reliable user authentication mechanisms

Forensics	Government	Commercial
Corpse identification	National ID card	ATM
Criminal investigation	Driver's license; voter registration	Access control; computer login
Parenthood determination	Welfare disbursement	Mobile phone
Missing children	Border crossing	E-commerce; Internet; banking; smart card

authentication techniques has increased in the wake of heightened concerns about security, and rapid advancements in networking, communication, and mobility. Thus, biometrics is being increasingly incorporated in several different applications. These applications can be categorized into three main groups (see Table 1):

1. Commercial applications such as computer network login, electronic data security, e-commerce, Internet access, ATM or credit card use, physical access control, mobile phone, PDA, medical records management, distance learning, etc.
2. Government applications such as national ID card, managing inmates in a correctional facility, driver's license, social security, welfare-disbursement, border control, passport control, etc.
3. Forensic applications such as corpse identification, criminal investigation, parenthood determination, etc.

Summary

The increased demand for reliable and convenient authentication schemes, availability of inexpensive computing resources, development of cheap biometric sensors, and advancements in signal processing have all contributed to the rapid deployment of biometric systems in establishments ranging from grocery stores to airports. The emergence of multibiometrics has further enhanced the matching performance of biometric systems [12, 13]. It is only a matter of time before biometrics integrates itself into the very fabric of society and impacts the way we conduct our daily business.

Related Entries

- ▶ Authentication
- ▶ Biometric Applications, Overview
- ▶ Enrollment
- ▶ Identification
- ▶ Soft Biometrics
- ▶ Verification

References

1. Jain, A.K., Flynn, P., Ross, A. (eds).: Handbook of Biometrics. Springer (2007)
2. Wayman, J.L., Jain, A.K., Maltoni, D., Maio, D. (eds.): Biometric Systems: Technology, Design and Performance Evaluation. Springer, New York (2005)
3. Bolle, R., Connell, J., Pankanti, S., Ratha, N., Senior, A.: Guide to Biometrics. Springer, New York (2003)
4. Wechsler, H.: Reliable Face Recognition Methods: System Design, Implementation and Evaluation. Springer (2006)
5. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003)
6. Jain, A.K., Ross, A., Prabhakar, S.: An Introduction to Biometric Recognition. IEEE Trans. Circuits Syst. Video Technol., Special Issue on Image- and Video-Based Biometrics **14**(1), 4–20 (2004)
7. Moon, H., Phillips, P.J.: Computational and Performance Aspects of PCA-based Face Recognition Algorithms. Perception **30**(5), 303–321 (2001)
8. Ratha, N.K., Connell, J.H., Bolle, R.M.: An analysis of minutiae matching strength. In: Proceedings of Third International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 223–228. Halmstad, Sweden (2001)
9. Prabhakar, S., Pankanti, S., Jain, A.K.: Biometric Recognition: Security and Privacy Concerns. IEEE Security Privacy Mag. **1**(2), 33–42 (2003)
10. Rejman-Greene, M.: Privacy issues in the application of biometrics: a european perspective. In: Wayman, J.L. Jain, A.K. Maltoni, D. Maio (eds.) D. Biometric Systems: Technology, Design and Performance Evaluation, pp. 335–359. Springer, New York (2005)
11. Kenny, S., Borking, J.J.: The Value of Privacy Engineering. J. Inform. Law Technol. (JILT) **7**(1), (2002)
12. Jain, A.K., Ross, A.: Multibiometric Systems. Commun. ACM, Special Issue on Multimodal Interfaces **47**(1), 34–40 (2004)
13. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. 1st ed. Springer, New York (2006)

BIP

► Biometric Technical Interface, Standardization

BIR

► Common Biometric Exchange Formats Framework Standardization

Blind Source Seperation

► Independent Component Analysis

Blood Vessel Wall

Blood vessel wall is the wall that forms the tubular channel of blood flow. The walls of arteries and veins consist of three layers, tunica intima, tunica media, and tunica externa (adventitia). The most inner layer, tunica intima consists of a surface layer of endothelium with a basement membrane. In the case of arteries, there is an internal elastic membrane around it. The tunica media is composed of smooth muscle and connective tissue. In the case of arteries, there is an external elastic membrane around it. The tunica externa is the connective tissue forming the outmost layer. The walls of capillaries consist of a single layer of endothelial cells and a basement membrane.

The cells of vessel walls control the function of blood vessels by receiving information not only from outside the wall but also from the crosstalk among them. For example, the smooth muscle of a vessel wall contracts or relaxes as the reaction to various agents. They include neurotransmitters, paracrine factors, hormones, and nitric oxide. By the constriction and the dilation, the inner diameter of the blood vessel changes and the blood pressure is controlled.

The inner surface of the vessel wall or the epithelial layer serves as a smooth barrier. Its selective permeability plays an important role in the balancing of fluid between blood and tissue fluid. The epithelial cells of the vessel wall also produce the various bioactive substances.

► Performance Evaluation, Overview

Brachycephalic

Brachycephalic is the head form characterized by an anteroposteriorly short and mediolaterally wide skull.

► Anatomy of Face

Branch-and-Bound Search

A search strategy employed by optimization algorithms in which the space of candidate solutions is navigated in a systematic manner via a tree-like structure by employing upper and lower bounds on the criterion function being optimized. The search strategy is characterized by three steps: (1) a branching step in which the space of possible candidate solutions is recursively partitioned, (2) a bounding step in which upper and lower bounds are estimated for the criterion function based on the candidate solutions within each partition, and (3) a pruning step in which improbable solutions are eliminated.

► Fusion, Feature-Level

Breeder Documents

Breeder documents are documents accepted for establishing ground truth with respect to an individual's identity. Breeder documents are typically used to obtain other identity documents or to establish/enroll an identity to obtain some benefit, privilege, or entitlement. One example of a breeder document is birth certificate, which may then be used to establish identity for the purpose of obtaining a driver's license.

► Fraud Reduction, Applications

C

Calibration

It refers to a software or hardware procedure to restore the initial operational conditions of a sensor or a device.

► Biometric Sensor and Device, Overview
► Image Formation

Camera

Camera is the face imaging device delivering either single image stills or video streams. It generally delivers a standard analog signal and requires a digitization device to produce digital images. Low-cost web cameras provide a standard USB interface to deliver digitized images in digital format. High-end cameras provide a standard digital signal for high-resolution images.

► Face Device
► Image Formation

Camera Device

► Face Device

Camera Model

The camera model describes how a point in the 3D space is projected on the 2D image plane. The projection that gives the 2D image coordinates of any point on the 3D space is defined, by the camera model and a few parameters (camera parameters) such, as the focal length.

► Face Pose Analysis

Camera Point of View

Camera point of view is the effective location and orientation of a camera that would result in the observed hand silhouette.

► Hand Data Interchange Format, Standardization

Cancelable Biometrics

Andy Adler
Carleton University, Ottawa, ON, Canada

Synonym

Revocable biometrics

Definition

Cancelable biometrics are designed to allow an individual to enroll and revoke a large number of different biometric samples. Each biometric image is encoded with a distortion scheme that varies for each application. The concept was developed to address the privacy and security concerns that biometric samples are limited and must be used for multiple applications. During enrollment, the input biometric image is subjected to a known distortion controlled by a set of

© 2009 Springer Science+Business Media, LLC

parameters. The distorted biometric sample can, in some schemes, be processed with standard biometrics algorithms, which are unaware that the features presented to them are distorted. During matching, the live biometric sample must be distorted with the same parameters, which must be securly stored. The cancelable nature of this scheme is provided by the distortion, in that it is not the user's "actual" biometric that is stored, but simply one of an arbitrarily large number of possible permutations. One concern with cancelable biometrics is the secure management of the distortion parameters.

Introduction

Cancelable biometrics describes a class of biometric matching algorithms designed to address the security and privacy concerns because of the limited number of biometric samples. This limitation – humans have only one face, two eyes and up to ten fingers – raises several concerns (► Security and Liveness, Overview): (1) the same biometric must be enrolled into multiple applications, potentially allowing cross application privacy and security vulnerabilities; for example, fingerprint images given to enter an amusement park may then be used to spoof a user to their bank; (2) a compromised biometric sample is a permanent loss to a user, unlike other security systems, where, for example, new cards or passwords can be issued; and (3) network protocols based on biometrics are potentially vulnerable to replay attacks.

Cancelable biometrics algorithms address these concerns by creating multiple varied independent biometric samples by processing the input image (or template features) with a parameterized distortion. The concept was developed by Ratha et al. [1] and subsequently extended by many others (e.g., [2, 3, 4, 5, 6]). During enrollment, the input biometric image is subjected to a known distortion controlled by a set of distortion parameters. The distorted biometric sample can, in some schemes, be processed with standard biometrics algorithms, which are unaware that the features presented to them are distorted. During matching, the live biometric sample must be distorted in exactly the same way, otherwise it cannot match the enrolled sample. This distortion must also satisfy the constraint that multiple different distortion profiles cannot match. Thus, the cancelable nature of this scheme is provided by the distortion, in that it is not the user's "actual" biometric, which is stored, but simply one of an arbitrarily large number of possible permutations. Cancelable biometrics is similar in some ways to biometric encryption (► Biometric Encryption), but differs primarily in that the goal of a cancelable biometric scheme is a *Match/Nonmatch* decision, while biometric encryption releases an encoded token or cryptographic key.

Two classes of cancelable biometrics are defined by Bolle et al. [2]: *signal* and *feature* domain distortions. For *signal domain distortion*, the raw biometric image is distorted. This image is subsequently processed by a traditional biometric system, which may be unaware of the distortions. Requirements for this scheme are that the distortion be large enough to create independent input images, but constrained such that the biometric system is able to identify and reliably register landmarks on the image. Examples are given for face, iris, and voice. For *feature domain distortion*, the biometric image is first processed to extract the template features, which are then distorted. This scheme may still use a traditional biometric template match algorithm. Feature distortion is recommended for fingerprints, where calculated minutiae are scrambled. It is difficult to envisage a simple fingerprint image distortion scheme, which destroys the original minutiae while still preserving a fingerprint-like image. It is emphasized that, much like a cryptographic hash function, there is no need to invert cancelable distortions, rather such distortions are designed to be one-way functions in which comparison is performed in the distorted space.

The first cancelable biometric algorithms for face recognition were proposed by [7]. The distortion takes place in the raw image space, since face recognition feature sets are not standardized. This places tight constraints on the nature of the distortion, since severely distorted faces will not be recognized and properly encoded by the algorithms. A different approach is taken by Savvides et al. [5] in which the cancelable distortion is tied to a face recognition algorithm based on correlation filters. Enrolled and test face images are distorted with a random kernel calculated from a key to generate an encrypted correlation filter. Since the same convolution kernel is present for both images, its effect is mathematically cancelled in the correlation filter. This scheme is somewhat similar

to the biometric encryption approach of Soutar et al. [9]. Boult [3] proposes a scheme in which face recognition feaures are encoded via scaling and rotation; The resulting data are separated into a "general wrapping" number, which is encrypted with a one-way transform, and a fractional part, which is preserved undestorted. Comparison is based on robust distance measures, which saturate at large distances.

The cancelable fingerprint templates of [8] use the minutiae rather than the raw image, since this allows both minutiae position and angle to be permuted (increasing the degrees of freedom of the transformation), and since distortion will interfere with the feature extraction process. The distortion is modeled on the electric field distribution for random charges. Results show a small impact on biometric errors (5% increase in FRR) over undistorted features. A theoretical approach to cancelable biometrics uses ▶ shielding functions [4], to allow a verifier to check the authenticity of a prover (user wanting to be verified) without learning any biometric information, using proposed δ-contracting and ϵ-revealing functions. The proposed system was based on simple Gaussian noise models and not tested with an actual biometric system. Unfortunately, it is unclear how practical functions can be found that account for the inherent biometric feature variability.

A "biohashing" approach has been proposed by Teoh et al. [6] and applied to many different modalities including fingerprint, face, and palm. This scheme applies a wavelet Fourier-Mellen transform (a rotation and scale invariant transform) to input images. Each bit of the tempate is calculated based on the inner product of the transformed image with a random image generated from a code. The claimed performance of this approach is 0% EER. Unfortunately, it has been shown by Kong et al. [10] that this high performance is actually due to the code being treated as a guaranteed secure password. Without this assumption, biohashing approaches show overall poor error rates.

In general, cancelable biometrics may be seen to represent a promising approach to address biometric security and privacy vulnerabilities. However, there are several concerns about the security of such schemes. First, there is very little work analyzing their security, except for an analysis of biohashing [10]. Secondly, while distortion schemes should be "preferably non-invertible" [2], no detailed proposed scheme has this property. In fact, it would appear to be trivial to "undistort" the template given knowledge of the distortion key in most cases. Third, cancelable biometrics would appear to be difficult to implement in the untrusted scenarios for which they are proposed: if the user does not trust the owner of the biometric sensor to keep the biometric private, how can they enforce privacy on the distortion parameters used? This last concern is perhaps the most serious: the security of cancelable biometrics depends on secure management of the distortion parameters, which must be used for enrollment and made available at matching. Furthermore, such keys may not be much better protected than current passwords and PINs. In summary, cancelable biometrics offer a possible solution to certain serious security and privacy concerns of biometric technology; however, current schemes leave a number of important issues unaddressed. Research is very active in this subject, and may succeed in addressing these concerns.

Related Entries

▶ Fingerprints Hashing
▶ Security and Liveness, Overview

References

1. Ratha, N., Connell, J., Bolle, R.: Cancelable biometrics. In Proc. Biometric Consortium Conference, Washington DC, USA (2000)
2. Bolle, R.M., Connell, J.H., Ratha, N.K.: Biometric perils and patches. Pattern Recogn. **35**, 2727–2738 (2002)
3. Boult, T.: Robust distance measures for face-recognition supporting revocable biometric tokens *Proc. 7th Int. Conf. on Automatic Face Gesture Recog* Southampton, UK, pp. 560–566 (2006)
4. Linnartz, J.-P., Tuyls, P.: New shielding functions to enhance privacy and prevent misuse of biometric templates. In Proc. AVBPA, Guiford, UK, LNCS **2688**, 393–402 (2003)
5. Savvides M, Vijaya Kumar BVK, Khosla, P.K.: Cancelable biometric filters for face recognition. In Proc. Int. Conf. Pattern Recognition, pp. 922–925 (2004)
6. Teoh, A.B., Ngo, D.C., Goh, A.: Biohashing: two factor authentication featuring fingerprint data and tokenised random number. Pattern Recogn. **37**, 2245–2255 (2004)
7. Ratha, N.K., Connell, J.H., Bolle, R.M.: Enhancing security and privacy in biometrics-based authentication systems. IBM Syst. J. **40**, 614–634 (2001)
8. Ratha, N., Connell, J., Bolle, R.M., Chikkerur, S.: Cancelable biometrics: A case study in fingerprints. Proc. Int. Conf. Pattern Recogn. **4**, 370–373 (2006)

9. Soutar, C., Roberge, D., Stoianov, A., Gilroy, R., Vijaya Kumar, B. V.K.: Biometric Encryption using image processing. Proc. SPIE Int. Soc. Opt. Eng. **3314**, 178–188 (1998)
10. Kong, A., Cheung, K.-H., Zhang, D., Kamel, M., You, J.: An analysis of biohashing and its variants. Pattern Recogn. **39**, 1359–1368 (2006)

Canonical Face Model

Canonical face model is the model that is used to store face images in databases. Once a facial image is acquired, it is resized and transformed to match the size and the orientation of the canonical face model, in which it is then stored in a database and used for face recognition tasks. For face recognition in documents, the canonical face model proposed by the International Civil Aviation Organization (ICAO) for the machine readable travel documents is used by many passport and immigration offices in many countries. This model stores faces using 60 pixels between the eyes, which ensures that feature-based face recognition techniques can be applied on these images. It has been argued however that this canonical face model may not be suitable for face recognition in video, due to the fact that face resolution in video is normally lower than 60 pixels between the eyes.

► Face Databases and Evaluation

CANPASS

► Iris Recognition at Airports and Border-Crossings
► Simplifying Passenger Travel Program

Capillary Blood Vessel

Capillary blood vessels are minute blood vessels, 5–10 μm in diameter, carrying blood from arterioles to venules. Blood flows from the heart to arteries. The arteries diverge into narrow arterioles, and then further diverge into capillaries. The blood in the capillaries is collected into venules. The venules converge into veins that carry the blood to the heart.

The capillaries are very thin and form a fine network called a capillary bed. The capillary wall consists only of a single layer of cells, the endothelium. This layer is so thin that it acts as a semipermeable membrane in the interchange of various molecules between blood and tissue fluid. It supplies oxygen, water, and lipids from the blood, and carries away the waste product such as carbon dioxide and urea from the tissue. Though each capillary is very thin, the capillaries form vast networks. The surface area of the capillary bed in the total body amounts to 6,000 m^2. Various stimulations from our living environment cause the change in the amount of blood in the capillary bed. In biometrics, this change should be taken into account particularly in venous pattern authentication.

► Performance Evaluation, Overview

Capture Volume

The volume in which an image-based biometric can be successfully captured.

► Iris Device
► Iris on the Move

Casts

A three-dimensional plaster cast can be made of a person's foot, as another aid in barefoot morphology comparison. The foot is placed in specially made foam, used in podiatry and orthotics, in order to capture the shape of the foot. Dental casting material can be poured into the foam, and, upon setting, a solid three-dimensional replica of the foot will be recovered.

► Forensic Barefoot Comparison

CBEFF

Common Biometric Exchange Formats Framework.

► Biometric Technical Interface, Standardization

CBEFF Biometric Data Block (BDB)

The BDB contains biometric data. The values of the mandatory CBEFF data elements, BDB Format Owner, and BDB Format Type encoded in the SBH identify the format of the BDB. A typical BDB could contain data conforming to one of the data interchange formats specified in ISO/IEC 19794 or one of the ANSI INCITS biometric data format standards (or a proprietary format).

► Common Biometric Exchange Formats Framework Standardization

CBEFF Biometric Information Records (BIRs)

BIRs are well-defined data structures that consist of two or three parts: the standard biometric header (SBH), the biometric data block (BDB), and the optional security block (SB). CBEFF permits considerable flexibility regarding BIR structures and BDB content, but does so in a way that makes it easy for biometric applications to evaluate their interest in processing a particular BIR.

► Common Biometric Exchange Formats Framework Standardization

CBEFF Security Block (SB)

The Security Block (SB) is an optional third component of Common Biometric Exchange Formats Framework Biometric Information Records (BIR). The SB may carry integrity data (e.g., digital signature or MAC (message authentication code)) or might also carry data associated with the encryption of the CBEFF Biometric Data Block (BDB). The format owner/format type approach was adopted to support the security block. This enables any public or private organization that wants to provide security solutions for BDBs and BIRs to identify and publish its security data formats in a standard way. The SB format owner/format type fields in the CBEFF Standard Biometric Header provide this SB identifier. CBEFF requires that if an integrity mechanism is applied to the BIR, then that mechanism must cover both the SBH and the BDB.

► Common Biometric Exchange Formats Framework Standardization

CBEFF Standard Biometric Header (SBH)

The header of a BIR (Standard Biometric Header – SBH) specifies metadata that describe specific characteristics of the biometric data contained in the data structures (e.g., biometric data format, modality, its creation date). It can also convey information useful to support security of the biometric data (e.g., security/integrity options), and other user-required data (e.g., user-defined payload, challenge-response data). CBEFF standards explicitly require that the SBH not be encrypted. This ensures that the header can always be examined by an application with the minimum necessary processing. CBEFF does, however, provide definitions for a couple of optional data elements that may be encrypted within the header.

► Common Biometric Exchange Formats Framework Standardization

CBEFF Wrapper

Synonym

BIR

Definition

The meta-data that is associated with a Biometric Data Block (BDB) to form a Biometric Information Record (BIR). The meta-data elements are specified in CBEFF Part 1, and the combination of these (from minimal to all) with one or more BDBs, for particular application areas, is specified in CBEFF Part 3.

▶ Biometric Technical Interface, Standardization

Central Retinal Artery and Vein

Main vascular trunks that supply (artery) and collect (vein) blood to and from the retina. Entering and exiting the retina at the optic disc they bifurcate across the retina forming its blood vessel network.

▶ Simultaneous Capture of Iris and Retina for Recognition

Cepstrum Transform

Cepstrum transform consists of the inverse Fourier transform of the logarithm of a signal in the frequency domain. Cepstral analysis has been widely used for separating signals from linear filtering. In this sense, if the speech signal is viewed as an output of a Linear Time-Invariant (LTI) system, where a source signal has passed through a filter, cepstrum transformation may be used to separate the source from the filter. As a homomorphic transformation, cepstrum presents a useful property, since the convolution of two signals can be expressed as the addition of their cepstra.

▶ Speaker Features

Chaff Points

Chaff points are additional fake minutiae used to hide the genuine minutiae, so that too many combinations exist for a brute force attack.

▶ Fingerprints Hashing

Challenge Response

An authentication mechanism in which the biometric system poses a question (challenge) to an individual and determines whether the latter provides a valid answer (response). This response may be used to validate the legitimacy of the biometric trait being presented to the system.

▶ Biometrics, Overview
▶ Keystroke Recognition

Charge Coupled Device (CCD)

For image processing, CCD is a type of sensor that utilizes motion of "buckets" of charge in response to electric fields. This is an older and more highly developed technology than CMOS image sensors. CMOS and CCD technologies are battling for dominance in the marketplace.

▶ Face Device
▶ Iris Device

Chrominance

In color images and videos, the chrominance (or shortly "chroma", the Greek word for color) are the components that contain the color information apart from the luminance.

▶ Skin Detection

Circular Hough Transform

The circular Hough transform detects circular features within an image. The image transformed is generated by computing gradients in the original image, and summing gradient values into each point that is a certain distance (the specified radius to search for) away from it. A circular edge within the original image will produce a peak value at the center of the circle in the image transform.

► Segmentation of Off-Axis Iris Images

Circumstantial Identification

Identification of victims based on circumstantial evidence, such as the victim's clothing, jewelry, and pocket contents.

► Dental Biometrics

Classification

► Supervised Learning

Classifier Cascade

In face detection, a classifier cascade is a degenerate decision tree where each node (decision stump) consists of a binary classifier.

In a classifier cascade, each node is a boosted classifier consisting of several weak classifiers. These boosted classifiers are constructed so that the ones near the root can be computed very efficiently at very high detection rate with acceptable false positive rate. Typically, most patches in a test image can be classified as faces/non-faces using simple classifiers near the root, and relatively few difficult ones need to be analyzed by nodes with deeper depth. With this cascade structure, the total computation of examining all scanned image patches can be reduced significantly.

► Face Detection

Classifier Combination

► Ensemble Learning
► Multiple Classifier Systems

Classifier Fusion

It is the main strategy used to combine classifier outputs in a multiple classifier system. In classifier fusion, each classifier contributes to the final decision for each input pattern.

► Multiple Classifier Systems

Classifier Selection

It is a strategy used to combine classifier outputs in a multiple classifier system. In classifier selection, each classifier is supposed to have a specific domain of competence (e.g., a region in the feature space) and is responsible for the classification of patterns in this domain.

► Multiple Classifier Systems

CLEAR

► Iris Recognition at Airports and Border-Crossings
► Registered Traveler

Client

A generic term for a person known by a biometric system, which grants individual privileges to its clients.

► Multiple Experts

Closed-Set Identification

Any subject presented to the biometric system for recognition is known to be enrolled in the system; thus, no rejection is needed in principle unless the quality of the input biometric trait is too low to process. It is the opposite of "Open-Set Identification."

► Performance Evaluation, Overview

CMOS Sensor

CMOS sensor is Complementary Metal Oxide Semiconductor. It is a solid state imaging sensor characterized by an integrated circuit containing an array of pixel sensors, each containing a photodetector and connecting to an active transistor reset and readout circuit.

► Face Device

Color Constancy

Color constancy is a feature of the human perception, where the perceived color of objects remains relatively constant under varying illumination conditions.

► Skin Detection

Commensurability

Commensurability is the property of being capable of direct comparison; literally "of common measure." The antithesis of commensurability is captured by the expression: "comparing apples with oranges." In biometrics, even within a single modality, commensurability issues arise, if biometric data takes the form of lists of features varying in length. How should a short list of features (e.g., minutiae extracted from one fingerprint) be compared with a longer list of features from another print? Should the excess (incommensurable) features from the longer list be considered as evidence of disagreement or simply as absence of evidence in the shorter list? An important design feature in iris recognition is that the IrisCode is always of fixed length (2,048 bits of data and 2,048 masking bits for the publicly deployed algorithm), regardless of how much or how little of an iris is visible and available for comparison with another iris. This commensurability greatly simplifies the matching process, and greatly accelerates its speed to about a million complete iris comparisons per second per 3GHz CPU, using simple Boolean string operators.

► Score Normalization Rules in Iris Recognition

Committee-Based Learning

► Ensemble Learning

Common Biometric Exchange Formats Framework Standardization

Fernando L. Podio[1], Fred Herr[2]
[1]National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA
[2]NIST Contractor - Identification Technology Partners, Inc, North Potomac, MD, USA

Synonyms

BIR; Biometric Information Record; SBH; Standard Biometric Header; SB; Security Block; Patron Format Specification; Biometric Registration Authority; Data Interchange Format

Definition

Common Biometric Exchange Formats Framework (CBEFF) provides a standardized set of definitions and procedures that support the interchange of biometric data in standard data structures called CBEFF biometric information records (BIRs). BIRs are well-defined data structures that consist of two or three parts: the standard biometric header (SBH), the biometric data block (BDB), and possibly the optional security block (SB). CBEFF permits considerable flexibility regarding BIR structures and BDB content, but does so in a way that makes it easy for biometric applications to evaluate their interest in processing a particular BIR. CBEFF imposes no restrictions on the contents of a BDB, which can conform to a standardized biometric data interchange format or can be completely proprietary. CBEFF standardizes a set of SBH data element definitions and their abstract values. A few of these data elements are mandatory in all SBHs (such as identifying the BDB format) and the rest are optional or conditional. Most of the data elements support description of various attributes of the BDB within the BIR. The optional SB provides a container for integrity and/or encryption-related data that must be available to validate or process the BIR and/or BDB (such as integrity signatures and encryption algorithm identity).

Introduction

At their conceptually simplest, standard CBEFF data structures promote interoperability of biometric-based application programs and systems by specifying a standardized wrapper for describing, at a high level, the format and certain attributes of the content of a biometric data record.

CBEFF data structures are called "Biometric Information Records (BIRs)". The header of a BIR (Standard Biometric Header – SBH) includes metadata that describe specific characteristics of the biometric data contained in the data structures (e.g., biometric data format, modality, and its creation date). The SBH can also convey information useful to support security of the biometric data (e.g., security/integrity options), and other user-required data (e.g., user-defined payload, challenge-response data). CBEFF standards explicitly require that the SBH not be encrypted (exclusive of, for example, channel encryption). This ensures that the header can always be examined by an application with the minimum necessary processing. CBEFF does, however, provide definitions for a couple of optional data elements that may be encrypted within the header.

The content of the Biometric Data Block (BDB) in a CBEFF BIR can be biometric data conforming to a biometric data interchange format standard or data that meet the requirements of a proprietary format (e.g., developed by vendors to support their own unique implementation features/processing). The BDB may be encrypted to protect the privacy of the data. Representative required abstract data elements defined in CBEFF standards for the SBH are the BDB format owner and type (which uniquely identify the format specification of the BDB) and BDB encryption/integrity options. A number of optional data elements are also specified such as the BDB biometric type (implicit in the BDB format), BDB creation date, and the validity period.

The optional third component of BIRs is the Security Block (SB). The SB may carry integrity-related data (e.g., digital signature or MAC (message authentication code) or might also carry data associated with the encryption of the BDB (e.g., key identification). The format owner/format type approach (used to indicate BDB format) was adopted to support the identification of the security block format. This enables any public or private organization that wants to provide security solutions for BDBs and BIRs to identify and

publish its security data formats in a standard way. The SB format owner/format type fields in the SBH provide this SB identifier. CBEFF requires that if an integrity mechanism is applied to the BIR, then that mechanism must cover both the SBH and the BDB.

CBEFF requires a *Biometric Registration Authority (RA)*. This RA has the responsibility to assign unique identifiers to biometric organizations. All biometric objects defined by the CBEFF standards (BDBs, Security Blocks, Products, Devices, Patron Formats) are uniquely identified by their 32-bit identifiers. The first 16 bits (the "owner" half of the field) are the identifier of the organization (assigned by the RA) that is responsible for the object. The second 16 bits (the "type") are assigned by the organization itself, which is responsible for maintaining whatever level of uniqueness required for its objects. The RA has the responsibility to publish the list of these identifiers where appropriate. The RA also publishes, if the owner desires, identifiers for objects that the owner wants to make available to the biometric community (for example, standards bodies have published the identifiers for their standardized patron formats and BDB formats; and some vendors have published the identifiers for some of their products). The CBEFF registry is located at http://www.ibia.org/cbeff/.

The format identifiers placed in the CBEFF SBH enable biometric applications to examine the SBH for the identifier values; if the application recognizes the value, it can then decide whether to process the biometric data in the BDB, but if it doesn't recognize the value, then it knows that it has not been designed to handle the particular form of data. At this time, the Registry can only be accessed by browser through the IBIA website; dynamic access from applications is not supported.

Every SBH is required to include the unique identification of its associated BDB format, expressed as the combination of the BDB Format Owner's identifier (which is a value assigned by the registrar) with the BDB Format Type identifier (which is a value assigned by the Format Owner, which can optionally register that value and provide access to the format specification through the Registry). This is the case with the two biometrics standards bodies, INCITS M1 (the InterNational Committee for Information Technology Standards – INCITS, Technical Committee M1 – *Biometrics*) and ISO/IEC JTC 1/SC 37 (ISO/IEC Joint Technical Committee 1 Subcommittee 37 – *Biometrics*), each of which has its own biometric organization value, and has registered several BDB format specifications (which are open standards available to the public). Conversely, biometric vendors who have developed their own proprietary data formats have, in some cases, registered those formats to make them available as widely as possible; but in other cases, have decided not to register them and only make them available to particular clients, partners, or customers.

CBEFF adds significant value in open and complex biometric systems, especially in cases where the system must cope with a wide variety of biometric data records, some of which may even be encrypted. The more easily decoded plain text of the CBEFF SBH is intended to greatly simplify the logic of the top levels of the system, which are responsible for routing each record to the correct biometric processing components. Equally important, where biometric data records are exchanged between different systems, the CBEFF SBH enables the interchange programs to do their work without ever having to "open" any of the records, since all the information they need to categorize and direct each record to its correct destination is in the plain text header. Some closed biometric systems (with no requirements for data interchange and interoperability with any other system) may not substantially benefit from the wrappers specified in CBEFF standards, especially in the cases where only one, or a very few, types of biometric data records (e.g., single biometric modality) may exist and where these records may be fairly quickly scanned to determine what biometric components should be called for processing.

CBEFF Patrons and Patron Formats

A **patron format** specification defines in full detail the structure of a particular BIR, including the actual encodings of the abstract values of the SBH fields. This includes the list of data elements that the format supports, how to locate each data element in the SBH, the values supported by each data element, and the correct encodings for each value. CBEFF is neutral regarding programming and encodings, leaving it to the patron to specify them as necessary in order to build successful patron format implementations. A patron format specification declares the patron's identifier for a specific patron format (this requirement is optional in the American National Standard INCITS 398 discussed in a later section). It should also include descriptive information about the intended use/environment of the format and any special

considerations for its use. Examples of patron format specifications are shown in Table 1.

In the CBEFF international standard (ISO/IEC 19785 addressed in a later section) CBEFF patrons are distinguished by their status as having open review and approval processes that ensure that their specifications follow the CBEFF standard's rules; are internally consistent; and will work in practice. As part of this vetting process, CBEFF requires that a patron format specification include a Patron Format Conformance Statement following a standardized form.

CBEFF Standards – Early Work

The initial version of CBEFF was developed by a technical development team formed as a result of three workshops sponsored by NIST and the Biometric Consortium, which were held in 1999. This version was published in January 2001 as NISTIR 6529 [1]. Further CBEFF development was undertaken under the umbrella of the Biometrics Interoperability, Performance, and Assurance Working Group cosponsored by NIST and the Biometric Consortium. In April 2004, an augmented and revised version of CBEFF was published as NISTIR 6529-A with a slightly modified title more accurately reflecting the scope of the specification [2]. In the meantime, in December 2002, the United States National Body, the American National Standards Institute, (ANSI) offered a draft version of NISTIR 6529-A as a contribution to JTC1/SC 37 – *Biometrics* for consideration as an international standard (JTC 1 is the Joint Technical Committee 1 of ISO/IEC). A new project for the development of an international version of CBEFF was approved in March 2003. In the U.S., NIST/BC offered the published version of NISTIR 6529-A to INCITS as a candidate American National Standards via fast track. The specification was published as ANSI INCITS 398–2005. ANSI INCITS 398–2005 contained the same text as NISTIR 6529-A.

CBEFF Standards – Recent and Current Work

Recent versions of the CBEFF standards have been developed by INCITS M1 and JTC1/SC 37, and the resulting standards are generally compatible with each other. In 2008 a revised version of ANSI INCITS 398–2005 was published as ANSI INCITS 398–2008 [3]. INCITS M1 is also developing a conformance testing methodology for CBEFF data structures specified in ANSI INCITS 398–2008.

JTC 1/SC 37 is responsible for the multi-part standard ISO/IEC 19785, Information technology — Common Biometric Exchange Formats Framework. Parts 1, 2 and 3 [4–6] are approved international standards, and Part 4 is progressing through its development stages. The sub-titles of the four parts are:

Part 1: Data element specification

Part 2: Procedures for the operation of the Biometric Registration Authority

Part 3: Patron Format Specifications

Part 4: Security block format specifications

Although ANSI INCITS 398 is a single part standard, its internal organization generally parallels that of ISO/IEC 19785. Each of these parts is described below.

ISO/IEC 19785 Part 1 (and the main clauses of ANSI INCITS 398):

This part of CBEFF defines the requirements for specifying the parts and structures of a BIR, as well as abstract data elements that are either mandatory in the BIR header or may optionally be included therein. Both standards define a BIR as having two required and one optional part: the standard biometric header (SBH), the biometric data block (BDB), and the optional security block (SB).

ISO/IEC 19785 Part 2:

The International Biometric Industry Association (IBIA) [7] has been performing the role of CBEFF RA for the CBEFF identifiers since the first CBEFF specification was published. ISO/IEC appointed IBIA as the RA for the international version of the standard. Part 2 defines in detail the RA responsibilities and procedures to be implemented by a Biometric Registration Authority to ensure uniqueness of CBEFF identifiers (i.e., patrons, format/product/security block owners, etc.). ANSI INCITS 398 does not replicate the equivalent level of detail, but still requires that the same registration authority be used to prevent ambiguity in identifying CBEFF objects.

Common Biometric Exchange Formats Framework Standardization. **Table 1** Patron format specifications

Patron format specifications published in ISO/IEC 19785 Part 3	
Clause 7: Minimum simple bit-oriented patron format	Encodes only mandatory abstract data elements from ISO/IEC 19785 Part 1. Specified in and uses ASN.1 PER-unaligned encoding rules. Does not support a Security Block.
Clause 8: Minimum simple byte-oriented patron format	Encodes only mandatory abstract data elements from ISO/IEC 19785 Part 1. Specified in 8 bit bytes, permitting any encoding mechanism that produces the required bit strings. Does not support a Security Block.
Clause 9: Fixed-length- fields, byte-oriented patron format using presence bit-map	Encodes mandatory and fixed-length-optional (but not variable length optional) abstract data elements. Encodes a bit map to indicate presence/absence of each optional data element in every instantiated SBH. Specified in 8 bit bytes, permitting any encoding mechanism that produces the required bit strings. Does not support a Security Block.
Clause 10: Fixed-length-fields, bit-oriented patron format using presence bit-map	Encodes, in the minimum possible number of bits, mandatory and fixed-length-optional (but not variable length optional) abstract data elements. Encodes a bit map to indicate presence/absence of each optional data element in every instantiated SBH. Specified in and uses ASN.1 PER-unaligned encoding rules. Supports a Security Block.
Clause 11: TLV-encoded patron format, for use with smartcards or other tokens	Specifies structure and content of an SBH for use with smartcards and similar technologies, taking advantage of their unique capabilities. Both byte-oriented and ASN.1 encodings are specified. Accounts for differences between on- and off-card matching requirements. Relies on the card's security mechanisms rather than using the CBEFF Security Block and encryption/integrity bits.
Clause 12: complex patron format	Similar to Clause 9, but supports all optional abstract data elements and supports multi-level BIRs. Byte-oriented specification and encoding. Supports a Security Block.
Clause 13: XML patron format	Supports all required and optional abstract data elements defined in Part 1. Provides both XML and ASN.1 schemas. Supports a Security Block.
Patron format specifications published in ANSI INCITS 398:2008	
Annex A: Patron Format A	Supports all abstract data elements defined in INCITS 398 clause 5, including a Security Block.
Annex B: Patron Format B	Supports the 3 abstract data elements required by a top-level structure in a multi-level BIR. In combination with Patron Format A, it is possible to encode multi-level BIRs having any number of levels.
Annex C: The BioAPI Biometric Identification Record (BIR)	Publishes, for convenience, the patron format specification from ANSI/INCITS 358–2002, Information Technology – The BioAPI Specification, 13 February 2002.
Annex D: ICAO LDS (TLV-encoded – for use with travel documents, smartcards, or other tokens)	Publishes, for convenience, the patron format specification developed by ICAO for machine readable travel documents (MRTDs). Note that the only similarity between this patron format and ISO/IEC 19785 Part 3, Clause 11 is that both are intended for smartcard environments but they are quite different in their content and structure.

Common Biometric Exchange Formats Framework Standardization. Table 1 (Continued)

Annex E: Patron Format PIV – NIST Personal Identity Verification (PIV)	Publishes, for convenience, the patron format specification required for applications conforming to the Personal Identity Verification (PIV) standard for Federal Employees and Contractors, Federal Information Processing Standard (FIPS) 201, and the associated NIST Special Publication 800–76–1 (SP 800–76–1), Biometric Data Specification for Personal Identity Verification.
Annex F: Patron Format ITL – NIST/ITL Type 99 Data Record	Publishes, for convenience, the patron format specification required in the law enforcement environment for the exchange of biometric data that is not supported by other logical records specified in the ANSI/NIST-ITL 1–2007 standard "Data Format for the Interchange of Fingerprint, Facial, & Other Biometric Information".

ISO/IEC 19785 Part 3:

Part 3 specifies several patron format specifications that conform to the requirements of Part 1. ANSI INCITS 398 also publishes several such specifications in annexes internal to the standard itself rather than in a separate part. There is no duplication of patron formats between the two standards; Table 1 below describes the patron formats included in each.

The BioAPI specification, ISO/IEC 19784–1 [8] publishes an important CBEFF patron format, the BioAPI BIR, in one of its annexes; this BioAPI BIR specification conforms to the 19785 Part 1 requirements. A standard application profile under development in JTC 1/SC 37 (Biometric Based Verification and Identification of Seafarers) [9] also specifies a CBEFF patron format (and security block format) for the Seafarer's ID (SID) document.

ISO/IEC 19785 Part 4:

This part of the standard is under development. Analogous to Part 3 and its specification of patron formats developed by JTC 1/SC37, the Part 4 draft standard is developing the specification for Security Block formats that support encryption of a BDB and integrity of a BIR. The application profile for Seafarers also specifies a CBEFF Security Block. The INCITS standard does not currently include any security block formats.

There are several minor differences between the ISO/IEC multi-standard and the INCITS standard.

1. The ISO/IEC standard relies on the application's implicit knowledge of its "domain of use" for determining the patron format specification and thus being able to parse the header. The patron formats specified by INCITS M1 include the patron format identifier in the SBH. This is a required feature for new formats that wish to conform to this standard (the requirement does not apply to other existing formats documented in the standard).
2. The ISO/IEC standard does not define the length or structure of abstract data elements of the SBH, but requires the patron format specification to provide the means for such determinations, which can in turn rely on encoding mechanisms (as in ASN.1 encoded records) or can specify other explicit means (e.g., inclusion of a length field). The INCITS standard explicitly defines abstract data elements for the lengths of each major structure in the SBH, but makes implementation of those data elements in the patron format specification conditional on whether some other means is provided (implicitly or explicitly) in the SBH. In practice, these requirements are equivalent.
3. The ISO/IEC standard defines five abstract data elements describing the entire BIR that parallel five elements that describe the BDB. This recognizes, for example, that the BIR's creation date may differ from the BDB's creation date if the BIR is assembled from BDB's retrieved from a database that was built earlier.

In practice these differences are indeed minor because both the ISO/IEC and INCITS standards define rules by which a patron format specification can specify additional SBH fields beyond the CBEFF abstract data elements. This provision ensures that patron format specifications are not prevented from addressing any special requirements they may have that are not anticipated by the standards.

CBEFF Flexibility and Adaptability

Structured as it is, with abstract data elements, a corresponding set of abstract values, and rules for their use defined in the base CBEFF standards (ANSI INCITS 398 and ISO/IEC 19785 Part 1), along with particular patron format specifications published as annexes in ANSI INCITS 398 and as Part 3 of ISO/IEC 19785, CBEFF supports – and demonstrates – great flexibility in satisfying unique requirements for data structures and contents. These standardized patron formats are useful in their own right, ranging from support of minimum requirements (in only 8 bytes) to complex BIRs containing many BDBs, each with its own SBH as part of a well defined structure. These formats also serve as examples of what the CBEFF data elements and rules for their use support in terms of the possible variations in patron formats.

Patrons may select a subset of the CBEFF data elements and values for a format specification, as long as they include those defined as mandatory by the standard. They may also impose stricter requirements on their users, such as making CBEFF-optional data elements mandatory in their new patron format or further constraining the range of values allowed. If the patron wants to support integrity and/or encryption in its environment then the specification must identify the mechanisms to be used and support any related data such as digital signatures or algorithm identifiers. Data elements for which CBEFF defines only a generic value can be restricted to very specific data content; conversely, if a CBEFF-defined data element "almost" satisfies a patron's requirements but would be better with more or different abstract values, then the patron is free to define those values in the patron format specification.

In addition to the standardized data elements and abstract values, CBEFF permits patrons to specify additional elements and values in support of unique or unanticipated requirements. These can be structural in nature to support decoding processes' navigation within the BIR, or they can be descriptive of attributes of the BDB that cannot be described by any of the CBEFF-defined elements. The CBEFF standard does require the patron to completely and unambiguously specify any such data element or value.

While the abstract level of CBEFF data elements and values is useful for the conceptual understanding of a CBEFF patron format, the careful specification of encoding requirements and syntax is critical to the successful implementation of interoperable biometric applications, especially where interchange of CBEFF BIRs between different biometrics-enabled systems is involved.

Here again the CBEFF standards permit virtually unlimited freedom for patrons to satisfy their unique requirements by developing format specifications tailored to their specific needs. The base CBEFF standards say almost nothing regarding data encoding, but they absolutely require any patron format specification to include detailed, unambiguous and complete encoding requirements for every aspect of the implemented BIRs. The patron formats in Table 1 provide correct examples of defining the encoding requirements of a patron format. Some of these use the various encoding rules of ASN.1; some define XML codes for the implementation; some are specified in a tabular format with each byte and bit specified as to its location and abstract meaning; and a couple use the tag-length-value (TLV) encoding for BIRs that are to reside on smart cards or other types of tokens.

Multiple BDBs in a BIR

Occasionally, a biometric system has a requirement to include more than one BDB in a single BIR. A system may need to keep one subject's BDBs of different modalities together or it may need to gather BDBs of a group of subjects into a single BIR. A legacy of the second version of CBEFF, NISTIR 6529A, is a set of data elements and syntax that supports concatenation and decoding of virtually any number of BDBs or complete BIRs into or out of a multi-layered single BIR. While this is quite workable for grouping a small number of BIRs, this approach does not provide support for finding and accessing a particular "simple" BIR within the collection.

ISO/IEC 19785 Part 3 (Clause 12) includes a patron format which defines the data elements and syntax for this structure. Neither of these approaches may be optimal for all applications. The CBEFF standards' multiple conceptual levels, from general abstractions to specific encoding requirements of individual patron formats, again provide the path to other solutions. Because CBEFF gives patrons the authority to define new abstract data elements, abstract values, data structures and the encodings to implement them, patrons

can specify BIR structures that meet their requirements for simplicity and efficiency. For example, direct access to any BDB in a multi-BDB BIR could be supported by a patron format that concatenates all the individual BIRs and then maintains pointers to each SBH and BDB in a top-level SBH that also contains suitable metadata about each included BIR. Using this approach, an application can efficiently process the top-level header to locate the single BIR it needs and then access it directly via the related pointers.

BIR Transformations

Both the ISO/IEC and ANSI INCITS versions of CBEFF recognize that there are situations where a BDB that is embedded in a CBEFF wrapper will be "transformed" into a wrapper of a different patron format (the BDB contents not being changed in any way). In this case, it is important that data elements describing attributes of the BDB content (such as BDB format and BDB creation date) carry the same information in the new BIR as in the old one, and CBEFF specifies rules to be followed for each CBEFF-defined data element. On the other hand, the information in some data elements may legitimately be different in the new BIR (such as BIR Creation Date and CBEFF Level). CBEFF specifies transformation rules that support the logical intent of the data element.

Conformance Testing Methodology Standards for CBEFF BIRs

INCITS Technical Committee M1 is developing a standard that addresses the requirements for testing conformance of instantiated BIRs to specific patron formats published within ANSI INCITS 398–2008. This draft standard specifies types of testing and test objectives, test assertions for particular patron formats, and test cases to implement the assertions. It is expected that when approved, the standard will include assertions and test cases for at least several of the ANSI INCITS 398 annexes.

Related Entries

► Biometric Technical Interface, Standardization
► Data Interchange Standards
► International Standardization of Biometrics, Overview

References

1. Podio, F.L., Dunn, J.S., Reinert, L., Tilton, C.J., O'Gorman, L, Collier, M.P., Jerde, M, Wirtz, B.: Common Biometric Exchange File Format. NISTIR 6529, January 2001
2. Podio, F.L., Dunn, J.S., Reinert, L., Tilton, C.J., Struif, B., Herr, F., Russell, J., Collier, M.P., Jerde, M., O'Gorman, L., Wirtz, B.: Common Biometric Exchange Formats Framework. NISTIR 6529-A, April 2004
3. ANSI INCITS 398–2008, American National Standard, for Information Technology –Common Biometric Exchange Formats Framework (CBEFF)
4. ISO/IEC 19785–1: 2006 Information technology - Common Biometric Exchange Formats Framework (CBEFF) – Part 1: Data element Specification – http://www.iso.org/iso/store.htm. Also adopted as INCITS/ISO/IEC 19785–1: 2006 2008. – http://webstore.ansi.org/
5. ISO/IEC 19785–2: 2006 Information technology – Common Biometric Exchange Formats Framework (CBEFF) – Part 2: Procedures for the operation of the Biometric Registration Authority – http://www.iso.org/iso/store.htm. Also adopted as INCITS/ISO/IEC 19785–2: 2006 2008. – http://webstore.ansi.org/
6. ISO/IEC 19785–3: 2007 Information technology – Common Biometric Exchange Formats Framework (CBEFF) – Part 3: Patron format specifications – http://www.iso.org/iso/store.htm. Also adopted as INCITS/ISO/IEC 19785–3: 2007 2008. - http://webstore.ansi.org/
7. International Biometric Industry Association CBEFF Registry: http://www.ibia.org/cbeff/
8. ISO/IEC 19784–1:2006 Information technology – Biometric application programming interface – Part 1: BioAPI specification – http://www.iso.org/iso/store.htm. Also adopted as INCITS/ISO/IEC 19784–1: 2006 [2007. – http://webstore.ansi.org/
9. JTC 1/SC 37 Final Committee Draft 24713–3, Biometric Profiles for Interoperability and Data Interchange – Part 3: Biometric Based Verification and Identification of Seafarers

Common Feature Approach

Common feature approach is a heterogeneous matching process in which the comparison is done between the templates of features common in both enrollment images and input probe images.

► Heterogeneous Face Biometrics

Comparison

► Palmprint Matching

Comparison Prints

These known exemplars of fingerprints are taken from donors under controlled conditions. The most common method used for recording these prints is ink on paper; however, this method is gradually being replaced by live scan technology. For law enforcement purposes, the prints of all 10 fingers are commonly recorded. For other purposes such as border control or passports, this may be limited to one or four fingers.

► Fingerprint Matching, Manual

Complementary Metal Oxide Semiconductor (CMOS)

CMOS is a widely used architecture for integrated circuits, particularly semiconductor memory circuits. CMOS has been adapted for use in image sensors and is a competitor to CCD image sensors. This is a newer sensor technology; its compatibility with the large scale integration techniques developed for semiconductor memory is a powerful advantage. CMOS and CCD technologies are battling for dominance in the marketplace.

► Iris Device

Compliance

► Conformance Testing for Biometric Data Interchange Formats, Standardization of

Computational Iris Recognition Systems

► Wavefront Coded® Iris Biometric Systems

Concept Drift

The target concept in a machine-learning task might change over time in terms of distribution, description, properties, etc. Often the changes are unpredictable, which brings problems to learning systems without self-adaptive mechanisms because the predictions of such systems might become less accurate as the time passes.

► Incremental Learning

Confidence Interval

A $100(1-\alpha)\%$ confidence interval for some parameter θ is a range of values (L, U) such that $P(\theta \epsilon (L, U)) = 1-\alpha$, where L and U are random variables.

► Test Sample and Size

Configural Processing

Configural processing is a specific mechanism that may evolve over time in which features are considered in relation or context with other features or areas. Configural processing typically allows more information to be extracted from each location or feature.

► Latent Fingerprint Experts

Configuration Issues, System Design

Kai Cao, Jie Tian, Yangyang Zhang, Xin Yang
Institute of Automation, Chinese Academy of Sciences, Beijing, People's Republic of China, China

Synonym

Setting

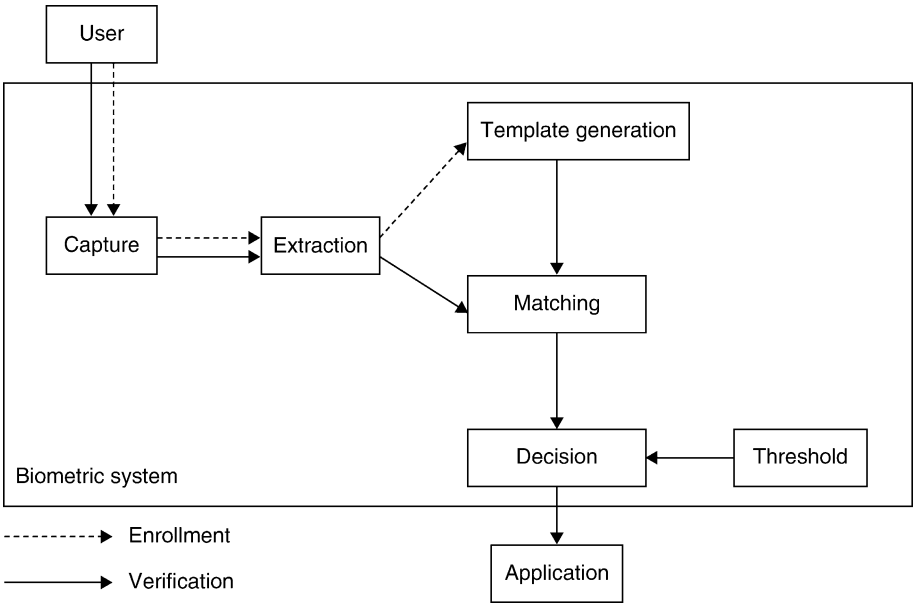
Definition

Configuration is to select appropriate parameters and/or fusion rules for an individual user or globally for all users of a system under a specific environmental condition (including the environment, the device, the received data, and the biometric algorithm).

Introduction

A typical biometric system can be divided into two distinct stages [1]: the enrollment stage which includes capture, feature extraction and template generation, and the verification or identification stage which comprises capture, feature extraction, matching and decision, as illustrated in Fig. 1. In the enrollment stage, for each user, a biometric trait is captured and processed to present it with a feature set. Then a template is generated and labeled with the user's identity. In the verification or identification stage the input biometric trait is processed as above and the system outputs the matching score or decision. Each component of the system in the two stages is critical and may be affected by the environmental condition. Care must be taken in any specific application since different kinds of applications focus on different requirements. Due to the variety of biometric system application in various environments and situations, configuration of selecting appropriate parameters and/or fusion rules to the biometric system is essential in order to satisfy different requirements.

For the component of capture, each device will have certain criteria to configure the capture process. For example, in a fingerprint device, the quality of the captured fingerprint image for enrollment should be high enough to ensure the reliable characteristic features (e.g., minutia) of the fingerprint. For face recognition devices, the person is usually required to be in a standard position directly facing the capture device. Threshold levels are also used in template generation process, where they determine the similarity required to make samples to be able to generate repeatable



Configuration Issues, System Design. Figure 1 A simplified verification system.

templates. All of these are unique considerations for biometric system.

Another important component is the decision process which mainly controls the security and accuracy. Security and accuracy is mainly controlled by the decision process. For a single biometric verifier the only parameter need to be determined is the threshold which can be configurable by an administrator or it may be fixed by the biometric system either for an individual user or globally for all users. If the matching score produced by the biometric system is larger than the threshold then the user is accepted, otherwise the user is rejected. However, the performance of the biometric system may be affected by the scenario of the application, environmental conditions, amount of user population, and many other factors. For example, iris recognition may depend on light levels, voice recognition may depend on ambient sound levels, and atmospheric dust levels may affect fingerprint devices. On the other hand, recognition based on any one of these modalities may not be sufficiently robust or else may not be acceptable to a particular user group or in a special situation because of restricted degrees of freedom and unacceptable error rates. One way to improve the performance is to install multiple sensors that capture different biometric traits, and this approach either focuses on a fixed set of traits or seeks greater flexibility through the implementation of systems which are more generically adaptable and reconfigurable. Configuration of selecting appropriate thresholds and fusion rules for this type of system increases complexity and raise questions, and is therefore regarded as a major challenge.

Configuration in Capture Process

Due to the changes in the environmental conditions (e.g., light level, dust, humidity, and cleanliness of the biometric capture device), the biometric system may capture biometric data of poor quality, which, especially in the enrollment stage, is responsible for many of the most matching errors in biometric systems so that it will limit the accuracy of the biometric system. For example, Tan et al. [2] pointed out, from the experiments on the ORL face database that the performance of an eigenface-based face recognizer drops quickly when the enrolled templates become poorly representative. The problem, however, may be alleviated by the control of the quality of the biometric data. If the input biometric data quality measurement is lower than a threshold configured by the system or the administrator, the device will reject the biometric data of poor quality and the capture process may be repeated. Quality level of the captured biometric data is expected to be high during the enrollment stage, since it forms the basis against which all further biometric matchings are made. Therefore, the quality threshold should be higher in the enrollment stage than that in the verification or identification stage. Some other criteria should also be considered. For example, if the area of captured fingerprint is too small it will impact false accept rate (FAR) and false reject rate (FRR).

From the consideration of security, the biometric system may be attacked by fake or artificial biometric traits. This case exists especially in fingerprint capturing. Fingerprint capture devices are deceived probably by well-duplicated fake fingers [3]. Therefore, it is necessary to detect fake biometric traits in order to ensure that only live biometric traits are capable of generating templates for enrollment and recognition. The biometric system may be configured to select approaches and parameters to ensure that the captured sample comes from a live human being.

Single-modal biometric system has limitations in terms of enrollment rates, and susceptibility to spoofing. A recent report [4] by the National Institute of Standards and Technology (NIST) to the United States Congress concluded that approximately two percent of the population does not have a legible fingerprint and therefore cannot be enrolled into a fingerprint biometrics system. However, this problem can be solved by employing multibiometrics in a layered approach. Therefore it is essential for the biometric system to provide an option to select a different biometric trait or employ a combination of these biometric traits to make the system adapt to different scenarios.

Configuration in Template Generation

As pointed out clearly by Uludag et al. [5], in real operational scenarios, input biometric data can exhibit

substantial variations compared to the templates collected during the enrollment stage of users. In other words, there may be a large intra-class variability, due to changes in the environmental conditions (e.g., illumination changes), aging of the biometric traits, variations of the interaction between the sensor and the individual (e.g., poor finger placement and facial expression of a person's face), all of which will limit the performance of the biometric system. In order to obtain good performance, configuration may be required to preserve distinct and repeatable biometric features from the user. This process is critical from a security evaluation point of view, since the level of uniqueness inherent in a template will influence FAR and FRR of the system. The first configuration parameter is the number of the templates, since the system accuracy can be improved by collecting multiple templates with different variation (e.g., different pose of a person's face) in the enrollment stage but the storage requirement will be huge if the number of the template becomes very big.

There are basically three possibilities of where to store the template: on a token, local unit, or on a central server. Each of these locations has its own advantages and disadvantages: (1) To store the template on a portable token such as smart card does not need to traverse the network. The user carries the template from location to location. The compare process will be very fast because only the template on the card needs to be considered. One drawback with this storage is that template on the smart card can be read by unauthorized individuals and the template will be stolen. (2) Storing on a central sever overcomes the problem of users authenticating from multiple locations. The input data will have to be transferred through the network. Therefore it provides an opportunity for a third party to intercept the data transfer and duplicate the biometrics data. The input data may be compared with every record data stored to identify the user. The speed will be very slow as the number of users grow. (3) The computation ability of local unit seems to be middle ground between central sever and small card. And distributively storing the data prevent a focal point of attack for malicious hackers. However, with this storage the security may be lacking because the template could be found on the hard drive.

On the other hand, performance of biometric recognition systems may degrade quickly when the time interval between the input biometric trait and the reference template is long because biometric features can change over time. The biometric system may be configured to update the reference template by re-enroll or during the matching operations, improve the reference template by merging and averaging minutiae of multiple biometric traits.

Configuration in Decision Process

Configuration in decision process for a biometric system can be classified coarsely into two categories in terms of biometric traits adopted by the system: single-modal biometric configuration and multi-modal biometric configuration. Both of these two classes of configuration are mainly to set the thresholds for each biometric device and/or the fusion rules algorithm among the sensors. The ability to estimate verifier error rates is very important for such configuration. Without knowledge of how well the system works under current configuration parameters, there is no way of knowing whether parameters should be changed and which parameter should be adjusted. Estimating verifier error rates provides feedback to remedy this situation.

Error Rate Estimation

There are mainly six important error rates [6] used to measure a biometric system: failure to acquire rate, failure to enroll rate, false match rate (FMR), false non-match rate (FNMR), false accept rate (FAR), and false reject rate (FRR). FAR and FRR are two of the most important indications of error rates of a biometric system. FAR is the expected proportion of transactions that a transaction will be erroneously accepted by the biometric system when it should have been rejected. FRR is the expected proportion of transactions that a transaction will be erroneously rejected when it should have been accepted.

From a security aspect, FAR has much more relevance than FRR. However, FRR can be considered as a measure of inconvenience but also a measure of availability, and needs to be kept within acceptable limits for the intended application. So, knowledge of FAR and FRR can be used to select appropriate configuration parameters. FRR is generally straightforward to determine, while FAR is often difficult to analyze in

operational system. FRR at the chosen decision threshold can be estimated by asking authentic users to report when they are rejected. This assumes that each time when an authentic user is rejected, he or she will report only once. Another way to estimate FRR is to represent transaction of authentic users to the system and record the probabilities of failure (or matching scores). The simplest way to estimate FAR and FRR simultaneously is to use stored transactions. Each transaction is compared with the other transaction of the same user and the probability of failure (or matching score) is record. All of the records form the genuine distribution. Similarly, the impostor distribution can be computed by comparing each transaction with transactions of other people using the stored transactions and recording the probabilities of failure (or matching scores). Then the obtained probability density (or matching score) curve can be used to estimate FRR and FAR at all threshold settings. All of these estimations should be done under the concrete environment, since biometric traits and performance of the system may be affected by their physical environmental condition. For example, a device to read iris patterns relies obviously on the ambient lighting conditions. This means that the physical environment needs to be defined as part of the biometric system's configuration, and also that tests should be conduct under the same or similar environmental condition.

Because of the expense of collecting large databases, performance of the devices or algorithms is usually measured on relative small databases. Wayman et al. [7] specifies four methods to evaluate FAR and FRR for large-scale identification system: (1) extrapolation from experiences; (2) identification as a succession of N verification; (3) extrapolation with extreme value; (4) extrapolation when the distance can be modeled. In fact, different environmental conditions such as illumination and humidity can also result in variation of FAR and FRR. Beattie et al. [8] proposed a structured approach based on expectation-maximization (EM) algorithm for estimating error rates.

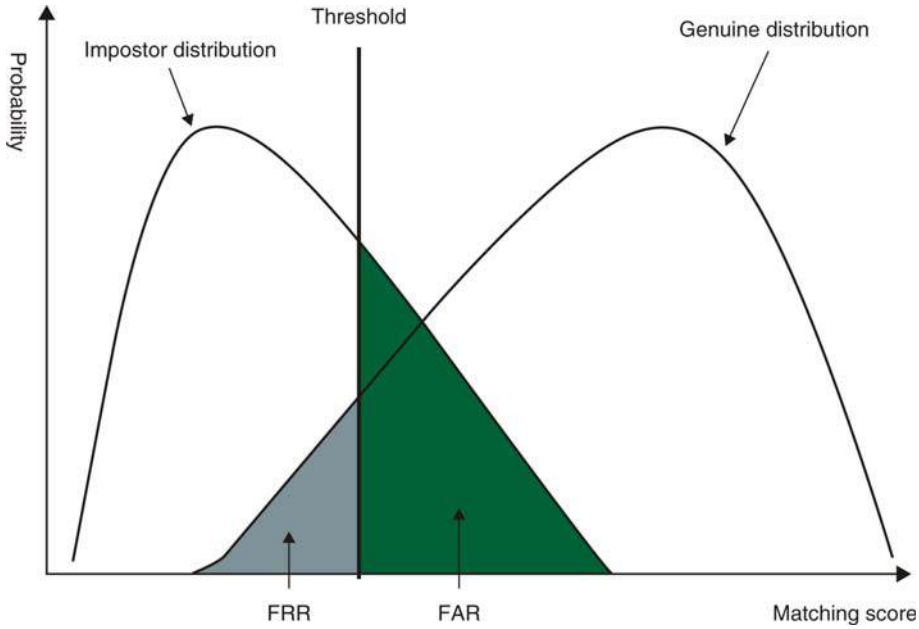
Single-Modal Biometric Configuration

In theory, matching scores of authentic users should always be higher than that of impostors. If this would be true, a single threshold that separates the two groups of scores could be used to differentiate between authentic users and impostors. Due to several reasons, this assumption isn't true for real-world biometric systems. In some cases, impostor patterns generate scores that are higher than the scores of some authentic patterns. For that reason, it is a fact that there is a threshold that needs to be determined to control the security and convenience of the biometric system. If the threshold is too high, the verifier will mostly reject decision regardless of whether the claimant is genuine or an impostor, the system is very secure but inconvenient. On the other hand, if the threshold is too low, it is convenient but insecure. Choosing a suitable threshold for a particular environment can be completed by analyzing the error rates on the training database captured in the same environment. The matching scores produced by the biometric system under the training database give a distribution for authentic users and another for impostors. When a threshold is set, FAR and FRR are simultaneously fixed, as illustrated in Fig. 2.

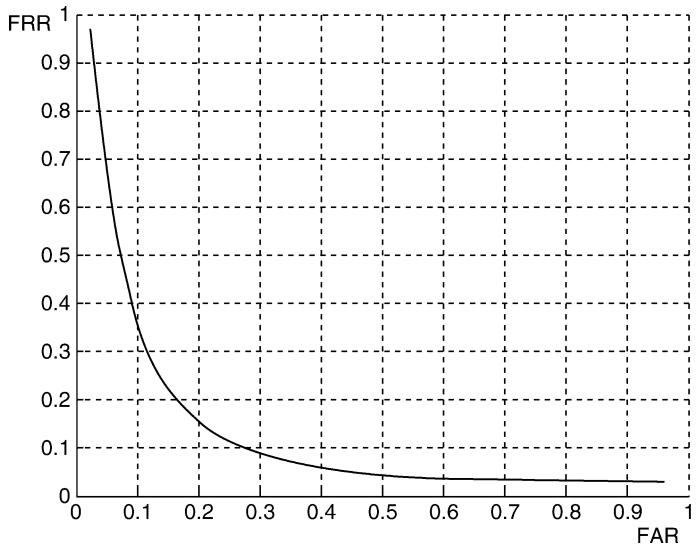
FAR and FRR are dependant on the adjustable threshold. If the value of threshold is increased, the proportion FAR will increase, while FRR will decrease. Otherwise, FAR will decrease and FRR will increase. This relationship between the two error rates is characterized by a receiver operating characteristic (ROC) curve in which FAR is plotted against FRR by varying the threshold, as shown in Fig. 3. Trade-off between FAR and FRR is often achieved by selecting an appropriate threshold so that the two rates can satisfy both the prescribed security and convenience.

The biometric system (e.g., physical access control) may be distributed over multiple locations where the environmental conditions may differ from each other. For this type of system, each biometric verifier must be configured with an appropriate threshold according to the performance under the environmental condition. In order to increase the reliability of single trait, the combination of different biometric matchers can be performed at the score level.

Combination of multiple matching algorithm is performed typically at the score level, and different fusion techniques (e.g., average, product, sum, max etc) have been applied successfully [7, 8]. The aim of multiple biometrics combined at the score level is to produce new scores whose distributions for genuine and impostor users exhibit a higher degree of separation than those produced by individual matchers. Thus, by varying the decision threshold, a better trade-off between FAR and FRR can be attained.



Configuration Issues, System Design. **Figure 2** FAR and FRR on normalized distributions graphs associated with a specific threshold.



Configuration Issues, System Design. **Figure 3** Receiver Operating Characteristic (ROC) curve.

Multi-Modal Biometric Configuration

In order to make the overlap of the distribution graphs of FAR and FRR in multi-modal biometric system, as little as possible, scores fusion rules can be adopted according to the confidentiality capture environment and recognition success rate of individual matchers. For this kind of multimodal biometric system not only the thresholds, but also the fusion rules need to determined. Knowledge of the characteristics of each biometric trait and sensor will be helpful to design an effective system. For example, from the experiment of

Ross et al. [9], the fingerprint and face biometrics generated a relatively small enroll failure rate, while the performance of the voice is much more stable than the other two modalities once a satisfactory enrollment has been achieved.

The most commonly adopted approach in multimodal biometric system is to use the same fusion rule and the same decision threshold for all users [9], the main idea of this system is to treat all matching scores from genuine users as one single class while all matching scores from imposter users as the other one. Ross et al. [9] combine the matching scores of three traits (face, fingerprint, and hand geometry) to enhance the performance of a biometric system. Experiments indicate that weighted sum rule outperforms other three techniques (sum rule, decision tree, and linear discriminate analysis) in terms of ROC curves. However, they do not mention how to configure the threshold. M. C. FairHurst [10] described an approach that used genetic algorithm (GA) to select appropriate parameters including weights and threshold to evolve efficient configuration using the Total Error Rate (FAR + FRR) of the overall system as an evaluation criterion.

Relatively, another new approach is using multiple fusion rules (each individual user correspond to a fusion rule or/and multiple decision thresholds (each individual user correspond to a threshold). Anil K. Jain [11] proposed a user-specific multimodal biometric system in which the common threshold is computed using the cumulative histogram of impostor matching score corresponding to each user and the user-specific weights associated with each biometric are selected by minimizing the total verification error. Toh et al. [12] improves this method using multivariate polynomial fusion model for each user. In other words, the system configures a personalized decision hyperplane (including thresholds) for each user.

Summary

Configuration plays an important role in the biometric system. A threshold in the decision process controls the trade-off between the security and the convenience. A multi-modal biometric system can utilize the predominance of each biometric trait and allow a more reliable biometric system. Accurate error estimation information would be useful to configure appropriate thresholds and/or fusion rules which will make the system more effective. Appropriate configuration will make the biometric system more robust, adaptive, and effective.

Related Entries

- ▶ Fusion, Score-level
- ▶ Multiple Classifier Systems
- ▶ Multi-modal systems
- ▶ Multibiometrics
- ▶ Performance Evaluation, Overview

References

1. Ross, A.A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. Springer, New York (2005)
2. Tan, X., Chen, S., Zhou, Z.-H., Zhang, F.: Face recognition from a single image per person: a survey. Pattern Recognit. **39**(9), 1725–1745 (2006)
3. Matsumoto, T.H., Yamada, K., Hoshino, S.: Impact of artificial 'gummy' fingers on fingerprint systems. In: Proceedings of SPIE, San Jose, CA, vol. 4677 (2002)
4. NIST report to the United State Congress. Summary of NIST standards for biometric accuracy, tamper resistance, and interoperability. http://www.itl.nist.gov/iad/894.03/NISTAPP_Nov02.pdf (2000). Accessed 13 Nov 2000
5 Umut, U., Ross, A., Jain, A.K.: Biometric template selection and update: a case study in fingerprints. Pattern Recognit. **37**(7), 1533–1542 (2004)
6. Common Criteria Biometric Evaluation Methodology Working Group - United Kingdom: Common Criteria - Common Methodology for Information Technology Security Evaluation - Biometric Evaluation Methodology Supplement. http://www.cesg.gov.uk/policy_technologies/biometrics/media/bem_10.pdf. Version 1.0, Aug., (2002)
7. Wayman, J., Jain, A., Maltoni, D., Maio, D.: Biometric Systems: Technology, Design and Performance Evaluation. Springer, New York (2005)
8. Beattie, M., Vijaya Kumar, B.V.K., Lucey, S., Tonguz, O.K.: Automatic configuration for a biometrics-based physical access control system. In: International Workshop on Biometric Recognition Systems (IWBRS), Beijing, 22–23 Oct, pp. 241–248 (2005)
9. Ross, A., Jain, A.K.: Information fusion in biometrics. In: Proceedings of International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), Sweden, pp. 354–359 (2001)
10. Fairhurst, M.C., Deravi, F., George, J.: Towards optimised implementations of multimodal biometric configurations. In: IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety, Venice, Italy, 21–22 July 2004

11. Jain, A.K., Ross, A.: Learning user-specific parameters in multibiometric system. In: Proceedings of International Conference on Image Processing (ICIP), Rochester, New York, USA pp. 57–60 (2002)
12. Toh, K.-Ann, Yau W.-Y.: Some learning issues in user-specific multimodal biometrics. In: Eighth International Conference on Control, Automation, Robotics and Vision Conference (ICARCV), Kunming, China, 6–9 Dec, vol. 2, pp. 1268–1273 (2004)

Conformance Testing

Conformance testing is the process of capturing the technical description of a specification and measuring whether an implementation faithfully implements the specification by achieving conformance to the technical description of the specification. Conformance is defined generally as the fulfillment by a product, process, or service of all relevant specified requirements.

► Conformance Testing for Biometric Data Interchange Formats, Standardization of

Conformance Testing for Biometric Data Interchange Formats, Standardization of

John W. M. Campbell[1], Gregory Zektser[2]
[1]Bion Biometrics Inc., ON, Canada
[2]US Department of Defense, Biometrics Task Force

Synonyms

Conformity; Compliance

Definition

The development of standardized methods and procedures for ► conformance testing of products or systems that claim to satisfy the requirements of one or more of the standardized biometric data interchange formats.

Concepts in Conformance Testing

A national or international standard consists of a set of requirements and frequently a set of recommendations. The requirements specified in the standard are traditionally classified in three categories: mandatory requirements to be observed in all cases, optional requirements to be observed if so chosen by an implementer, and conditional requirements to be observed under specific conditions. A product, process, or system that fully satisfies the requirements of the standard is described as being conformant to that standard. Conformance testing is the method that is used to determine if the product, process, or system satisfies the requirements. The precise nature of these requirements differs substantially from one standard to another, but in all cases, there are certain important concepts.

The product, process, or system being tested is known as an ► implementation under test or IUT. It does not need to satisfy every requirement and recommendation of a standard, only those that that are defined as mandatory. In certain cases, the mandatory requirements may be different for different applications or purposes. In the case of conformance testing for biometric data interchange format standards, for example, an IUT may be designed to produce biometric data interchange records or to use biometric data interchange records or both. In each case, the requirements that are tested for conformance would be different.

No conformance test can be complete or perfect. Ultimately, it is only possible to prove that an implementation under test is nonconformant. The goal of conformance testing is therefore to capture enough of the requirements of the base standard and test them under enough conditions, that any IUT that passes the conformance test is likely to be conformant. Frequently, there are inherent problems with the underlying standards that only become apparent during conformance testing. For instance some areas may be undefined (so that the specification of these areas is left to each vendor) or ill-defined (so that there is a contradiction between parts of the base standard or an easy misinterpretation caused by the wording of the base standard). The latter problem may be resolved by an amendment to or revision of the standard, but the former problem may be difficult to resolve.

Conformance testing does not guarantee interoperability; it can only provide a higher level of confidence that interoperability can be achieved. Although the

ultimate goal of standards is to allow different products and systems to work together, even two products that are conformant to the same standard may have difficulty working together. This is because it is usually impossible for a standard to specify every aspect of the operation of a product or system. On the other hand, unless systems are conformant to a well written standard, then it is very unlikely that they will be interoperable. Thus conformance testing is a critical element in assuring interoperability, even if it is not the only one.

Motivation for the Development of Conformance Testing Methodology Standards

As increasing numbers of biometric standards have been developed in recent years, more and more products have become available that claim to be conformant to the standards. This is particularly true in the area of standardized biometric data interchange formats which are standard methods of encoding biometric data for various technologies, including 2D face, 3D face, finger image, finger pattern, finger minutiae, iris image, vein pattern, signature/sign, hand geometry, etc. Theoretically, those products that support the relevant standard for a given technology should be able to work together, so that an end user of biometrics can mix products from a variety of vendors or support interoperability among different systems.

Although vendors provide products and systems in good faith believing that they conform to a standard, if there is no corresponding conformance testing methodology standard, then there is no clear method for them to be able to verify this. Similarly, end users of biometric products cannot know with confidence if the products and systems they are using actually conform to the standards unless a formal conformance testing methodology standard exists and can be used to perform conformance testing on those products and systems in a reliable and repeatable manner.

Elements Required in Conformance Testing Methodology Standards for Data Interchange Formats

In order to formally describe conformance testing for data interchange formats, it is necessary to identify a language to define the context of conformance testing and conformance claims. Therefore a number of specialized terms have been developed. Many of them relate to the fact that there are lots of different types of testing that can be defined for different levels and types of conformance. In the standardization process that has taken place in the US at INCITS M1 – Biometrics and internationally at ISO/IEC JTC 1/SC 37 – Biometrics, the following key elements have been defined.

Test Assertion – The specification for testing a conformance requirement in an IUT in the forms defined in a conformance testing methodology standard. Test assertions are short specific statements that encapsulate a single requirement for a particular standard. A harmonized assertion description language has been developed for data interchange format conformance testing so that the assertions can be expressed in a simple form, regardless of the specific data interchange format being addressed.

Level 1 Testing – A set of test methods within the conformance testing methodology that validates field by field and byte by byte conformance against the specification of the ▶ Biometric Data Interchange Record as specified in the base standard, both in terms of fields included and the ranges of the values in those fields.

Level 2 Testing – A set of test methods within the conformance testing methodology that tests the internal consistency of the Biometric Data Interchange Record (BDIR) under test, relating values from one part or field of the BDIR to values from other parts or fields of the BDIR.

Level 3 Testing – A set of test methods within the conformance testing methodology that tests that a Biometric Data Interchange Record produced by an IUT is a faithful reproduction of the input biometric data that was provided to the IUT.

Type A – Produce Conformant BDIR (Type A or PCB) – A conformance claim by an IUT that it is a conformant BDIR, or can create conformant BDIRs from an appropriate input data.

Type B – Use Conformant BDIR (Type B or UCB) – A conformance claim by an IUT that it can read conformant BDIRs, interpret them correctly, and perform its desired function upon them.

Issues Related to Testing Levels

It is obvious from the carefully defined terminology listed above that there are issues that have led the

standardization bodies to separate the different levels and types of testing. The main consideration is the need for a balance between the importance of delivering conformance testing methodology standards that are meaningful and that can be used to support testing and the desire to thoroughly test all aspects of each data interchange format standard.

The first issue is the fact that data interchange format standards are mostly focused on the structure and content of the BDIR. This means that the test assertions for Level 1 testing can be simply developed by analyzing the explicit requirements of the standard. Test assertions for Level 2 testing may require consideration of the implicit requirements of the standard, but they can still be defined quite specifically. Some experts prefer to state that Level 1 testing supports the syntactic requirements of the standard and Level 2 testing supports the semantic requirements of the standard. Unfortunately, some semantic requirements can only be addressed through Level 3 testing, and because of the inherently uncertain nature of biometric data, it is very difficult to establish a standardized method of determining whether a BDIR is or is not a faithful reproduction of the input biometric data used to produce it. Human biometric characteristics vary with every presentation to a biometric system and there is debate among experts on exactly how to define the relationship between the BDIR and the input characteristic, especially when it comes to acceptable levels of accuracy in the representation. For this reason, Level 3 testing is still an area of research and has not been included in the conformance testing standards that are currently published or under development.

The second issue relates to the fact that the BDIR itself is the focus of the biometric data interchange format standards. It is therefore easy to test claims of Type A conformance, since the output BDIRs can be tested at least for Level 1 and Level 2 conformance. An IUT that claims Type B conformance, however, needs to interpret the BDIRs correctly and perform its appropriate function upon them. Since this function may be to use them for biometric matching, to display them for human examination, to convert them to another format or potentially a whole host of other things, it is very difficult to determine how best to test such claims of conformance. One option is to force IUTs to also support specific functions of usage that would only be used in Type B conformance testing, but so far this idea has not been popular among biometric vendors or standardization experts. It remains to be seen how Type B conformance testing will be addressed in the future.

Conformance Testing Standardization – Current State

The need for standardized and commonly accepted conformance testing methodologies for Biometric data interchange formats has been recognized by the National and International Standards Bodies on Biometrics. In February 2005, INCITS M1 initiated the development of a multi-part American National Standard on conformance testing methodology for Biometric Data Interchange Formats. This project is based on an extensive analysis of the data format requirements specified in the base data interchange format standards, and is structured to take advantage of the commonalities found in the testable requirements as well as in the conformance test methods and procedures. The resulting structure of this multi-part standard is as follows:

1. Part 1: Generalized Conformance Testing Methodology
2. Part N: Modality-specific Testing Methodology (e.g., Part 2: Conformance Testing Methodology for Finger Minutiae Data Interchange Format)

The Generalized Conformance Testing Methodology contains the elements of the testing methodology that are common to all the data interchange formats (i.e., those elements that are modality independent). These elements include definitions of terms, descriptions of levels and types of testing, general requirements of test reports, specification of the assertion definition language, general test procedures, etc.

Each individual Part contains elements of the testing methodology specific to its respective modality. These elements include specific definitions of terms, specifications of test assertions, test criteria, modality-specific elements of test reports, test procedures, etc.

At the time of preparation of this paper, the Part 1 of this multi-part standard has been published as INCITS 423.1, and other parts are in various stages of development ranging from publication stage to Public Review stage. It is expected that most of the Parts of this standard will be publicly available in 2008.

While the development of the INCITS 423 was underway, Working Group 3 of ISO/IEC JTC 1/SC

37 initiated the development of a similar multi-part international standard in 2006. This ISO/IEC Project 29109, named "Conformance Testing Methodology for Biometric Data Interchange Formats defined in ISO/IEC 19794" is similarly structured, and also consists of Part 1: Generalized Conformance Testing Methodology, and multiple modality-specific Parts, each dedicated to one modality. At the time of preparation of this paper, Part 1 is being circulated to the National Bodies of JTC 1/SC 37 for a Committee Draft (CD) ballot, Parts 2 (Finger Minutiae), 4 (Finger Image), 5 (Face Image), 6 (Iris Image), and 10 (Hand Geometry) are in the Working Draft (WD) stage, and a number of other Parts are expected to be presented at the July 2008 SC 37 meeting.

Conformance Testing Activities

Approval and publishing of the conformance testing methodology standards alone does not ensure conformance of the Biometric products to the base standards. It is imperative that the published testing standards are adopted by the Biometric community, including technology vendors, system integrators, and end-users, and implemented in the form of conformance testing tools, processes, and programs. Some of these efforts are already underway, although at the time of publication of this paper there are very few large-scale conformance testing and conformity assessment/certification programs for Biometric data interchange formats.

The fact that a number of Biometric industry vendors claim conformance of their products to national and international data interchange format standards suggests that at least some first-party conformance testing (vendor self-testing) is taking place. It is not known whether the standardized conformance testing methods and procedures are used for this testing.

There are indications that governments are interested in establishing second- or third-party conformance testing programs. For example, the United States Department of Defense described their Biometric Conformity Assessment Initiative in [1] that includes the standards-based conformance testing and reporting of Biometric products, although it is not known when this program will be fully implemented.

Two large scale conformance testing programs have been established ahead of the publication of the necessary standards, and the methods used in these programs have influenced the development of the standards. In the US, the certification for biometric algorithms to be approved for use with personal identity verification (PIV) associated with Homeland Security Presidential Directive 12 (HSPD-12) requires that they be tested in a program called MINEX. This testing ensures that biometric templates produced by the template generation algorithms are conformant to a profiled version of INCITS 378:2004 – Finger Minutiae Format defined specifically for PIV [2]. Similarly, template generation algorithms that are part of biometric products to be used with the Seafarers' Identity Documents programme associated with the International Labour Organization Convention No. 185 [3] must be tested by a third party laboratory and found to be conformant to a profiled version of ISO/IEC 19794-2:2005 – Finger Minutiae Data.

Current and Anticipated Needs

It is reasonably well understood that the major needs in implementations of the Biometric systems can be described as interoperability of the systems on all levels and ability to interchange the Biometric data. These needs can be fulfilled, to a significant extent, by standardization of all aspects of Biometric technology, including Biometric formats for data interchange. Such standardization requires the following:

1. Robust base standards must exist and be commonly accepted
2. Biometric technology must be implemented in conformance with the base standards
3. End-users must be able to verify conformance of the implementation to the standards

The last element by itself can be further decomposed in to the following:

1. Standardized conformance testing methodologies must exist and be commonly accepted
2. Conformance testing tools implementing the standardized methodologies must exist
3. Laboratories performing the conformance testing must exist and be able to produce standardized test results

4. A process of certification of test results by an independent authority must exist

As shown above, development of the conformance testing methodology standards is only the first necessary step in establishing the conformance testing programs that would be able to reliably test Biometric products and provide reasonably conclusive determination of conformance (or nonconformance) of the products to the base standards. While publishing of the conformance testing methodology standards, currently under development, and expeditious development of conformance testing tools that implement these standards is recognized as an immediate need, establishing of such full-scale conformity assessment programs in the near future should be anticipated.

Gaps in Standards Development

The development of the conformance testing methodology standards in national and international standards development bodies is progressing quite rapidly, and it is not unreasonable to expect completion of majority of these development projects within the next 24 months. There are, however, certain gaps in the existing projects that will need to be addressed at some point in the future, for the testing methodologies to remain useful. These gaps can be divided into three categories:

1. *Completeness of the standard.* Currently, the conformance testing methodology standards don't provide (and probably will never provide) full, absolute coverage of all requirements of the base standards. For example, Type B and Level 3 testing are currently out of scope of the existing Parts of the conformance testing methodology standards. The motivation behind this is based on practical reasons, and on the fact that certain requirements can not be tested in a reasonable manner; nonetheless the conformance testing coverage is not 100% conclusive. It is expected that additional test cases/assertions will be developed as the conformance testing methodologies mature, but it is unlikely that the desirable full coverage will ever be reached.
2. *Coverage of modalities.* Currently, even if most of the existing modalities' conformance testing standardization is planned, many of the Parts have not been initiated even as preliminary drafts. For some relatively new modalities, such as DNA or Voice, it is not even clear how conformance testing should be performed. It is fully expected that eventually conformance testing methodologies will be developed for all modalities, but at the present time this is a significant gap.
3. *The testing methodologies are almost always "behind" the base standards.* The base standards, however robust and mature, are always undergoing changes, amendments and revisions. These changes, sometimes significant, may not be immediately be reflected in the corresponding conformance testing standard, and the time gap between the base standard change and the conformance testing methodology standard corresponding change may be significant – from several months to several years.

Summary

The increased need for interoperability of Biometric systems, especially their ability to interchange and share biometric data records has driven the demand for standardization of nearly every aspect of the Biometric technology. One of the primary elements of this standardization effort has been development of the Biometric Data Interchange Format Standards and corresponding conformance testing methodologies that ensure fulfillment by the biometric implementations of the requirements specified in the base standards.

References

1. Woodward, J., Cava, S.: DoD biometric conformity assessment initiative. Defense Standardization Program J., April/June 2005
2. Wilson, C., Grother, P., Chandramouli, R.: Biometric Data Specification for Personal Identity Verification, NIST Special Publication 800-76-1 (2007)
3. Seafarers' Identity Documents Convention (Revised), International Labour Organization Convention No. 185, International Labour Organization (2003)

Conformity

► Conformance Testing for Biometric Data Interchange Formats, Standardization of

Contact Microphones

Contact microphones are special microphones that transduce not sound, but vibrations in solid bodies into electrical signals. These can be used, for instance, to capture speech directly from the throat's surface, which is an interesting possibility in very noisy environments.

► Voice Device

Contact-Based

It refers to a device for which it is needed to touch the sensing area to image of the ridge-valley pattern.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Contactless

Manner of biometric recognition in which it is not necessary for the person to touch the sensor. Some people are reluctant to touch a publicly used device for various reasons, and some contact-type sensors tend to become dirty through use. There are high expectations that the contactless-type authentication method can be used in public situations and will prove valuable in places that require a high level of hygiene such as hospitals.

► Palm Vein

Contextual Biases

Contextual biases are external influences on a perceptual decision due to outside factors such as details from a case.

► Latent Fingerprint Experts

Continuous Classification

► Fingerprint Indexing

Contour Detection

Contour detection is the coarse edge detection technique that extracts the outer boundary of an object, without extracting inner edges.

► Hand Geometry

Contrast

In general terms, contrast is a measure of the difference between two objects, concepts, or other entities. In photography, optics, and image acquisition contrast usually refers to the difference in hue, saturation, or intensity between two portions of an image.

► Iris Device
► Photography for Face Image Data

Convenience Sample

A convenience sample is a sample that uses individuals or sample units that are readily available rather than those that are selected to be representative or selected via a probabilistic mechanism.

► Test Sample and Size

Convergence Feature Extraction

Convergence provides a more general description of channels and wells than force field feature extraction. It takes the form of a mathematical function in which wells and channels are revealed to be peaks and ridges, respectively, in the function value. This function maps the force field $\mathbf{F}(\mathbf{r})$ to a scalar field $C(\mathbf{r})$, taking the force as input, and returning the additive inverse of the divergence of the force direction, and is defined by,

$$C(\mathrm{r}) = -\mathrm{div}\,\mathbf{f}(\mathbf{r}) = -\lim_{\Delta A \to 0} \frac{\oint \mathrm{f}(\mathbf{r}) \cdot d\mathbf{l}}{\Delta A} = -\nabla \cdot \mathbf{f}(\mathbf{r}) = -\left(\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y}\right) \quad (1)$$

where $\mathbf{f}(\mathbf{r}) = \frac{\mathbf{F}(\mathbf{r})}{|\mathbf{F}(\mathbf{r})|}$ is the force direction, ΔA is incremental area, and $d\mathbf{l}$ is its boundary outward normal. This function is real valued and takes negative values as well as positive ones, where negative values correspond to force direction divergence. Note that the function is non-linear because it is based on force direction and therefore must be calculated in the given order.

► Physical Analogies for Ear Recognition

Copula

A copula is a multivariate joint distribution that is defined on the n-dimensional unit cube $[0; 1]^n$ such that every marginal distribution is uniform on the interval $[0; 1]$. A copula can be used to capture the dependencies or associations that exist between variables in a multivariate distribution. In the context of score-level fusion, a copula may be used to describe the correlation between multiple matchers.

► Fusion, Score-Level

Core

The topmost point on the innermost recurving ridge of a fingerprint. Generally, the core is placed upon or within the innermost recurve of a loop as described in the Standards document ISO/IEC 19794-2: Biometric Data Interchange Formats – Part 2: Finger Minutiae Data.

► Fingerprint Templates

Correct Index Power

The ratio of correctly retrieved fingerprints over the size of the database.

► Fingerprint Indexing

Correct Reject Power

The ratio of correctly rejected reference fingerprints over the number of query images not having a corresponding fingerprint in the database.

► Fingerprint Indexing

Correlation

The degree of relationship between two variables as expressed using a single measure. The Pearson correlation coefficient is an example of one such measure. In the context of multibiometrics, the correlation between the genuine (or impostor) match scores of two biometric matchers can have a bearing on the performance of the fusion scheme used to combine them.

► Multibiometrics

Correlation Map

Correlation map is a two dimensional array of correlation values in the range $[-1, +1]$. These correlation values are obtained by computing the normalized local correlation between two curvature feature maps. Each pixel in the correlation map represents local correlation between the corresponding pixels in the curvature maps being matched.

► Palmprint Features

Correspondence

► Human Detection and Tracking

Cost Function

Tracking is an estimation process that computes the position of a target based on optimization of a criterion that relates the observations with the estimates. The criterion is represented mathematically using the cost function.

► Face Tracking

Countermeasures

Liveness Check is a validation that the biometric characteristic is the true characteristic of the presenting individual by the measurement of expected live features such as pulse, temperature, humidity, movement etc. as appropriate to the biometric characteristic.

Artifact detection is the detection of an artifact that has been presented by measurement of specific property of known artifacts (e.g., silicone rubber or gelatine finger; photograph of face etc. depending on the biometric characteristic). A point to be noted here is that the liveness check and artifact detection are complementary approaches to countering the use of artifacts.

Biometric data encryption is a cryptographic technique used to safeguard the confidentiality of biometric data.

Biometric data signing is also a cryptographic technique used to safeguard the integrity of biometric data.

Cryptographic timestamps/session keys are cryptographic techniques used to counter capture/replay attacks.

Supervised operation is a powerful countermeasure against a range of threats that can occur when a subject is interacting with a biometric system during enrolment and verification operations. It can be an effective countermeasure to the use of artifacts, mimicry, and physical attacks.

Security audit is a useful post-event analysis of security log to check e.g., suspicious events, integrity of system configuration, procedural compromises etc.

Performance audit is an offline check that the system performs to a level that safeguards security, e.g., cross checking enrolment references against each other to ensure that adequate separation between references exists and there are no apparent cases of multiple enrolments by a single individual.

► Biometric Security, Standardization

Counter Sign

1. A second sign to provide the proofs of approval and/or receipt as on a previously signed document such as a contract or a money order.

2. Another sign, word, or signal used for replying the sign from an anonymous or a hidden person.

► Signature Matching

Covariate

A covariate is a secondary variable or factor that can affect the relationship between the dependent variable and other independent variables of primary interest. For instance, in biometrics view-point or illumination is a secondary variable that can impact the relationship between recognition ability, the dependent variable, and identity, the independent variable. Covariate needs to be controlled or monitored in a biometric experiment.

► Evaluation of Gait Recognition

Covariate Studies

In multivariate statistical analysis the aim is to study the independence and interdependence of two or more random variables. In such a study, there could be other confounding factors that affect the statistical analysis of the variables of interest. The study of these confounding variables and design of appropriate theoretical and experimental set-ups in order to eliminate the effect of these confounding random variables is called covariate analysis.

For instance, in designing statistical methods for face recognition, the statistical relationship between the identity variable and the face images obtained has to be studied. The face images obtained are also dependent on various confounding factors such as illumination, pose, expression, and age of the subject, and the camera internal parameters. These confounding variables are considered covariates for the problem of image/video-based face recognition. Similarly, in the case of gait-based person identification from videos, it would be interesting to learn the statistical relationship between the subject's identity and the videos. In this case, there are several environmental confounding variables such as the clothing of the subject, the shoe-type of the subject, the surface of walking, the camera view and the presence of other occluding objects enter a briefcase etc.

Two major approaches are available deal with to the effect of covariates – enumeration and marginalization. Enumeration refers to techniques where one learns the statistical relationship between the random variables of interest for every possible realization of the confounding variable. This approach works well for settings such as the presence/absence of a briefcase, where there are very few distinct values that the confounding random variable can take. In marginalization a joint probability density function of the relevant and the confounding random variables is first developed and then marginalized (integrated) over the confounding random variables in order to make an inference.

► Gait Biometrics, Overview

Craniofacial Reconstruction

► Skull, Forensic Evidence of

Craniofacial Superimposition

► Skull, Forensic Evidence of

Credential Hardening

The process of increasing the trust associated with typical credential sets such as user names and passwords is credential hardening. This increased trust is often achieved through a biometric augment as in keystroke recognition. This is sometimes also referred to as password hardening.

► Keystroke Recognition

Credentialing System

Credentialing System uses a physical credential, such as a smartcard, as a means to authenticate the identity of a credential holder for purposes of authorization. It includes the registration process as well as the subsequent operational use of the credential. Registration may include enrollment, identity proofing, background checking, card production, and issuance. Possible uses are many, but generally include physical and logical access control as well as benefits eligibility/redemption and other privilege or entitlement claims.

▶ Registered Traveler

Credit Check

▶ Background Checks

Crew Designs

▶ Test Sample and Size

Criminal History Check

▶ Background Checks

Criminal Law Enforcement

▶ Law Enforcement

Criminal Record Search

▶ Background Checks

Cross-Modality Face Biometrics

▶ Heterogeneous Face Biometrics

Cross-Validation

A popular approach to estimating how well the result learned from a given training data set is going to generalize on unseen new data. It partitions the training data set into k subsets with equal size, and then uses the union of k_i1 subsets for training and the remaining subset for performance evaluation. The final estimate is obtained by averaging after every subset has been used for evaluation once. A popular setting of k is 10 and in this case it is called as *10-fold cross-validation*; another popular setting of k is the number of training examples and in this case it is called as LOO (i.e., *Leave-One-Out*) test.

▶ Ensemble Learning

Cryptography

The science of transforming messages or data into incomprehensible formats for the purposes of confidentiality, integrity, authentication, or non-repudiation of origin. Cryptographic systems have classically involved two parties, a sender and a receiver whom wish to communicate a message secretly, although modern uses include secure data storage as well as digital signatures. Despite the intended use, cryptographic systems, sometimes referred to as cryptosystems involve two main blocks: encryption and decryption. Encryption is the process of encoding data into an unreadable format

(sometimes referred to as ciphertext) through the use of a cryptographic key. Decryption is the process of decoding encrypted data (ciphertext) back into a comprehensible format through the use of a cryptographic key and a complementary algorithm tied to encryption. Cryptographic algorithms can be symmetric, requiring the communication of a single secret key between sender and receiver or they can be asymmetric relying on public-private key pairs that do not explicitly require the transmission between sender and receiver.

► Iris Digital Watermarking

Curse of Dimensionality

The demand for a large number of samples grows exponentially with the dimensionality of the data (feature) space, and so is the difficulty to find global optima for the parameter space, i.e., to describe the data space. This phenomenon is known as the "Curse of Dimensionality." The fundamental reason for this limitation is that high-dimensional functions have the potential to be much more complicated than low-dimensional ones, and that those complications are harder to discern. A simple but effective way to alleviate this problem is to reduce the number of dimensions of the data by eliminating some coordinates that seem irrelevant or extract and select salient and discriminatory features for data representation.

► Biometric Algorithms
► Fusion, Feature-Level
► Multibiometrics

Curse of Misalignment

► Face Misalignment Problem

Cursive

In a fully cursive handwriting style, the writer is inclined to connect all adjacent letters in a smooth way. What makes it more complex is that different people may have different habits to connect even the same character pair. For example, some people like to connect the t-bar with other letters, while others may connect the t-stem to neighboring characters. Real-world handwriting is often in a mixture of cursive and handprint styles.

► Signature Sample Synthesis

Custody Suite

Custody suites are areas of police stations in which suspects arrested in connection with particular circumstances are held and questioned in the furtherance of the enquiries by the police.

► Footwear Recognition

Cut Finger Problem

► Anti-Spoofing
► Fingerprint Fake Detection
► Liveness Detection

D

Dactyloscopist: Fingerprint Examiner

► Fingerprint, Forensic Evidence of

Data Hiding

Data hiding techniques can be used to insert additional information, namely the watermark, into a digital object. The watermark can be used for a variety of applications ranging from copy protection, to data authentication, fingerprinting, broadcast monitoring, multimedia indexing, content based retrieval applications, medical imaging applications, etc. Within the framework of biometrics, robust data hiding techniques can be used to embed codes or timestamps into the template, in such a way that after the expiration date the template is useless. Another perspective is to hide the biometric template in a digital object to make it invisible when either transmitted or stored.

► Conformance Testing for Biometric Data Interchange Formats, Standardization of
► Iris Template Protection

Data Interchange Format

► Common Biometric Exchange Formats Framework Standardization

Data Protection

► Privacy Issues

Database Filtering

Filtering refers to limiting the number of entries in a database to be searched, based on characteristics of the interacting user. For example, if the user can be identified as a middle-aged male, the search can be restricted only to the subjects with this profile enrolled in the database. This greatly improves the speed or the search efficiency of the biometric system. Filtering reduces the probability of obtaining a wrong match, but this is offset by the fact that the errors in filtering also reduce the probability of obtaining a correct match. Hence, in general, filtering drastically reduces the time required for identification but can degrade the recognition accuracy.

► Soft Biometrics

Daubert Standard

The Daubert standard is a legal precedent set in 1993 by the Supreme Court of the United States regarding the admissibility of expert witnesses' testimony during

© 2009 Springer Science+Business Media, LLC

federal legal proceedings (the citation is Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579).

► Gait, Forensic Evidence of

Daugman Algorithm

► Iris Encoding and Recognition using Gabor Wavelets

Dead Finger Detection

► Fingerprint Fake Detection

Decision

Decision is the output result of a biometric system given a biometric sample. For a verification task, the decision is "reject or accept" of the claimed identity, while for identification task it is the identity of the presented subject or rejecting him or her as one of the enrolled subject.

► Performance Evaluation, Overview

Decision Criterion Adjustment

► Score Normalization Rules in Iris Recognition

Deformable Models

Thomas Albrecht, Marcel Lüthi, Thomas Vetter
Computer Science Department, University of Basel, Switzerland

Synonyms

Statistical Models; PCA (Principal Component Analysis); Active (Contour, Shape, Appearance) Models; Morphable Models

Definition

The term Deformable Model describes a group of computer algorithms and techniques widely used in computer vision today. They all share the common characteristic that they *model* the variability of a certain class of objects. In biometrics this could be the class of all faces, hands, or eyes, etc. Today, different representations of the object classes are commonly used. Earlier algorithms modeled shape variations only. The shape, represented as curve or surface, is *deformed* to match a specific example in the object class. Later, the representations were extended to model texture variations in the object classes as well as imaging factors such as perspective projection and illumination effects. For biometrics, deformable models are used for image analysis such as face recognition, image segmentation, or classification. The image analysis is performed by fitting the deformable model to a novel image, thereby parametrizing the novel image in terms of the known model.

Introduction

Deformable models denote a class of methods that provide an abstract model of an object class [1] by modeling separately the variability in shape, texture, or imaging conditions of the objects in the class. In its most basic form, deformable models represent the shape of objects as a flexible 2D curve or a 3D surface that can be deformed to match a particular instance of that object class. The deformation a model can undergo is not arbtitrary, but should satisfy some problem-specific constraints. These constraints reflect the prior

knowledge about the object class to be modeled. The key considerations are the way curves or surfaces are represented and the different form of prior knowledge to be incorporated. The different ways of representing the curves range from parametrized curves in 2D images, as in the first successful method introduced as Snakes in 1988 [2], to 3D surface meshes in one of the most sophisticated approaches, the 3D Morphable Model (3DMM) [3], introduced in 1999. In the case of Snakes, the requirement on the deformation is that the final deformed curve should be smooth. In the 3DMM, statistical information about the object class (e.g., such as the class of all faces) is used as prior knowledge. In other words, the constraint states that the deformed surface should with high probability belong to a valid instance of the object class that is modeled. The required probability distributions are usually derived from a set of representative examples of the class.

All algorithms for matching a deformation model to a given data set are defined as an energy minimization problem. Some measure of how well the deformed model matches the data has to be minimized. We call this the *external energy* that pushes the model to match the data set as good as possible. At the same time the *internal energy*, representing the prior knowledge, has to be kept as low as possible. The internal energy models the object's resistance to be pushed by the external force into directions not coherent with the prior knowledge. The optimal solution constitutes an equilibrium of internal and external forces. For instance, in the case of Snakes, this means that a contour is pushed to an image feature by the external force while the contour itself exhibits resistance to be deformed into a non-smooth curve. In the case of the 3DMM, the internal forces become strong when the object is deformed such that it does not belong to the modeled object class.

This concept can be expressed in a formal framework. In each of the algorithms, a model $\mathcal{M}$ has to be deformed in order to best match a data set $\mathcal{D}$. The optimally matched model $\mathcal{M}^*$ is sought as the minimum of the energy functional E, which is comprised of the external and internal energies E_{ext} and E_{int}:

$$E[\mathcal{M}] = E_{\text{ext}}[\mathcal{M}, \mathcal{D}] + E_{\text{int}}[\mathcal{M}] \tag{1}$$

$$\mathcal{M}^* = \arg\min_{\mathcal{M}} E[\mathcal{M}]. \tag{2}$$

Snakes

Kaas et al. introduced Snakes, also known as the Active Contour Model in their landmark paper [2]. Here, the deformable model $\mathcal{M}$ is a parametrized curve and the goal is to segment objects in an image $\mathcal{D}$ by fitting the curve to object boundaries in the image. The external energy $E_{\text{ext}}[\mathcal{M}, \mathcal{D}]$ measures how well the snake matches the boundaries in the image. It is expressed in form of a feature image, for instance, an edge image. If an edge image I with low values on the edges of the image is used, the external energy is given as:

$$E_{\text{ext}}[\mathcal{M}, \mathcal{D}] = E_{\text{ext}}[v, I] = \int_0^1 I(v(s))\, ds, \tag{3}$$

where $v : [0,1] \rightarrow \mathrm{IR}^2$ is a suitable parametrization of the curve $\mathcal{M}$ and $I : \mathrm{IR}^2 \rightarrow \mathrm{IR}$ is the edge image of the input image $\mathcal{D}$. If a point $v(s)$ of the curve lies on a boundary, the value of the edge image $I(v(s))$ at this point is low. Therefore, the external energy is minimized if the curve comes to lie completely on a boundary of an image.

The internal energy ensures that the curve always remains a smooth curve. For the classical snakes formulation, it is defined as the spline bending energy of the curve:

$$E_{\text{int}}[\mathcal{M}] = E_{\text{int}}[v] = (\alpha(s)|v'(s)|^2 + \beta(s)|v''(s)|^2)/2, \tag{4}$$

where α and β control the weight of the first and second derivative terms.

By finding a minimum of the combined functional $E[\mathcal{M}]$, the aim is to find a smooth curve $\mathcal{M}$, which matches the edges of the image and thereby segments the objects present in the image.

The Snake methodology is the foundation for a large number of methods based on the same framework. There are three main lines of development:

- Flexible representation of curves and surfaces
- Incorporation of problem specific prior knowledge from examples of the same object class
- Use of texture to complement the shape information

Level Set Representation for Curves and Surfaces

The idea of snakes was to represent the curve $\mathcal{M}$ as a parametric curve. While such a representation is simple, it is topologically rigid, i.e., it cannot represent objects that are comprised of a variable number of independent parts. Caselles et al. [4] proposed to represent the curve $\mathcal{M}$ as a level set, i.e., the contour is represented as the zero level set of an auxiliary function ϕ:

$$\mathcal{M} = \{\phi = 0\}. \tag{5}$$

A typical choice for ϕ is the distance function to the model $\mathcal{M}$.

This representation offers more topological flexibility, because contours represented by level sets can break apart or join without the need of reparametrization. Additionally, the level set formulation allows a treatment of surfaces and images in any dimension, without the need of reformulating the methods or algorithms. The idea of representing a surface by a level-set has led to a powerful framework for image segmentation, which is referred to as *level-set segmentation*.

Example Based Shape Priors

Before the introduction of Active Shape Models [5], the internal energy or prior knowledge of the Deformable Model has been very generic. Independent of the object class under consideration, the only prior knowledge imposed was a smoothness constraint on the deformed model. Active Shape Models or "Smart Snakes" and the 3DMM [3] incorporate more specific prior knowledge about the object class by learning the typical shapes of the class.

The main idea of these methods is to assume that all shapes in the object class are distributed according to a *multivariate normal distribution*. Let a representative training set of shapes $\mathcal{M}_1, \ldots, \mathcal{M}_m$, all belonging to the same object class be given. Each shape $\mathcal{M}_i$ is represented by a vector x_i containing the coordinates of a set of points. For 2D points (x_j, y_j), such a vector x takes the form $x = (x_1, y_1, \ldots, x_n, y_n)$. For the resulting example vectors $x_1, \ldots, x_m$, we can estimate the mean $\overline{x}$ and covariance matrix Σ. Thus, the shapes are assumed to be distributed according to the multivariate normal distribution $\mathcal{N}(\overline{x}, \Sigma)$. To conveniently handle this normal distribution, its main modes of variation, which are the eigenvectors of Σ, are calculated via ▶ Principal Components Analysis (PCA) [6]. The corresponding eigenvalues measure the observed variance in the direction of an eigenvector. Only the first k most significant eigenvectors $v_1, \ldots, v_k$ corresponding to the largest eigenvalues are used, and each shape is modeled as:

$$x = \overline{x} + \sum_{i=1}^{k} \alpha_i v_i, \tag{6}$$

with $\alpha_i \in \mathrm{IR}$. In this way, the prior knowledge about the object class, represented by the estimated normal distribution $\mathcal{N}(\overline{x}, \Sigma)$, is used to define the internal energy. Indeed, looking at Equation (6), we see that the shape can only be deformed by the principal modes of variation of the training examples.

Furthermore, the coefficients α_i are usually constrained, such that deformations in the direction of v_i are not much larger than those observed in the training data. For the Active Shape Model, this is achieved by introducing a threshold $D_{\max}$ on the mean squares of the coefficients α_i, scaled by the corresponding standard deviation σ_i of the training data. The internal force of the Active Shape Model is given by:

$$\begin{aligned} E_{\mathrm{int}}[\mathcal{M}] &= E_{\mathrm{int}}[\alpha_1, \ldots, \alpha_k] \\ &= \begin{cases} 0 & \text{if } \sum_{i=1}^{k} \left(\frac{\alpha_i}{\sigma_i}\right)^2 \le D_{\max} \\ \infty & \text{else.} \end{cases} \end{aligned} \tag{7}$$

In contrast, the 3DMM [3] does not strictly constrain the size of these coefficients. Rather, the assumed multivariate normal distributions $\mathcal{N}(\overline{x}, \Sigma)$ is used to model the internal energy of a deformed model $\mathcal{M}$ as the probability of observing this model in the normally distributed object class:

$$\begin{aligned} E_{\mathrm{int}}[\mathcal{M}] &= E_{\mathrm{int}}[\alpha] = -\ln P(\alpha) \\ &= -\ln e^{-\frac{1}{2}\sum_{i=1}^{k} (\alpha_i/\sigma_i)^2} \\ &= \frac{1}{2}\sum_{i=1}^{k} (\alpha_i/\sigma_i)^2. \end{aligned} \tag{8}$$

Correspondence and Registration

All deformable models using prior knowledge in form of statistical information presented here assume the

example data sets to be *in correspondence*. All objects are labeled by the same number of points and corresponding points always label the same part of the object. For instance in a shape model of a hand, a given point could always label the tip of the index finger in all the examples. Without this correspondence assumption, the resulting statistics would not capture the variability of features of the object but only the deviations of the coordinates of the sampled points. The task of bringing a set of examples of the same object class *into correspondence* is known as the *Registration Problem* and constitutes another large group of algorithms in computer vision.

Incorporating Texture Information

One limitation of the classical Snake model is that the information of the data set $\mathcal{D}$ is only evaluated at contour points of the model $\mathcal{M}$. In level-set ▶ segmentation, new external energy terms have been introduced in [7, 8]. Instead of measuring the goodness of fit only by the values of the curve $\mathcal{M}$ on a feature image, in these new approaches the distance between the original image and an approximation defined by the segmentation is calculated. Typical approximations are images with constant or smoothly varying values on the segments. This amounts to incorporating the prior knowledge that the appearance or texture of the shape outlined by the deformable model is constant or smooth.

By incorporating more specific prior knowledge about the object class under consideration, the appearance or texture can be modeled much more precisely. This can be done in a similar fashion to the shape modeling described in the previous section. The appearance or texture $\mathcal{T}$ of a model $\mathcal{M}$ is represented by a vector T. All such vectors belonging to a specific object class are assumed to be normally distributed. For instance, it is assumed that the texture images of all faces can be modeled by a multivariate normal distribution. Similar to the shapes, these texture vectors need to be *in correspondence* in order to permit a meaningful statistical analysis.

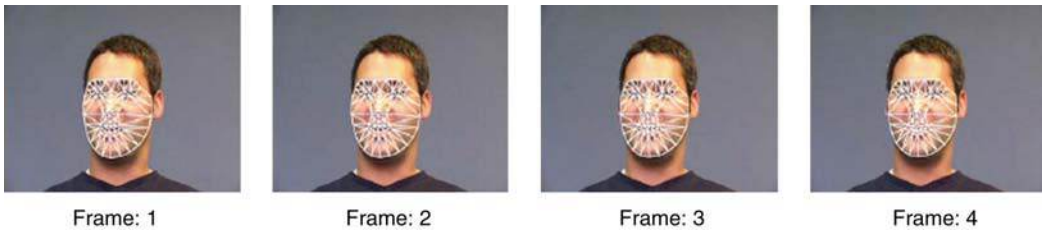
Given m example textures $T_1, \ldots, T_m$, which are in correspondence, their mean $\overline{T}$, covariance matrix Σ_T, main modes of variation $t_1, \ldots, t_k$, and eigenvalues ρ_i can be calculated. Thus, the multivariate normal distribution $\mathcal{N}(\overline{T}, \Sigma_T)$ can be used to model all textures of the object class, which are then represented as:

$$T = \overline{T} + \sum_{i=1}^{k} \beta_i t_i. \tag{9}$$

A constraint on the coefficients β_i analogous to Equation (7) or (8) is used to ensure that the model texture stays in the range of the example textures. In this way, not only the outline or shape of an object from the object class but also its appearance or texture can be modeled. The Active Appearance Models [1, 9, 10] and the 3D Morphable Model [3] both use a combined model of shape and texture to model a specific object class. A complete object is modeled as a shape given by Equation (6) with texture given by Equation (9). The model's shape and texture are deformed by choosing the shape and texture coefficients $\alpha = (\alpha_1, \ldots, \alpha_k)$ and $\beta = (\beta_1, \ldots, \beta_k)$. The external energy of the model is defined by the distance between the input data set $\mathcal{D}$ and the modeled object (S, T), measured with a distance measure which not only takes the difference in shape but also that in texture into account. The internal energy is given by Equation (7) or (8) and the analogous equation for the β_i.

2D versus 3D Representation

While the mathematical formalism describing all previously introduced models is independent of the dimensionality of the data, historically the Active Contour, Shape, and Appearance Models were only used on 2D images, whereas the 3DMM was the first model to model an object class in 3D. The main difference between 2D and 3D modeling is in the expressive power and the difficulty of building the deformable models. Deformable models, when incorporating prior knowledge on the objects class, are derived from a set of examples of this class. In the 2D case these examples are usually registered images showing different instances of the class. Similarly, 3D models require registered 3D examples. As an additional difficulty, 3D examples can only be obtained with a complex scanning technology, e.g., CT, MRI, laser, or structured light scanners. Additionally, when applied to images the 3D models require a detailed model for the imaging process such as the simulation of occlusions, perspective, or the effects of variable illumination.



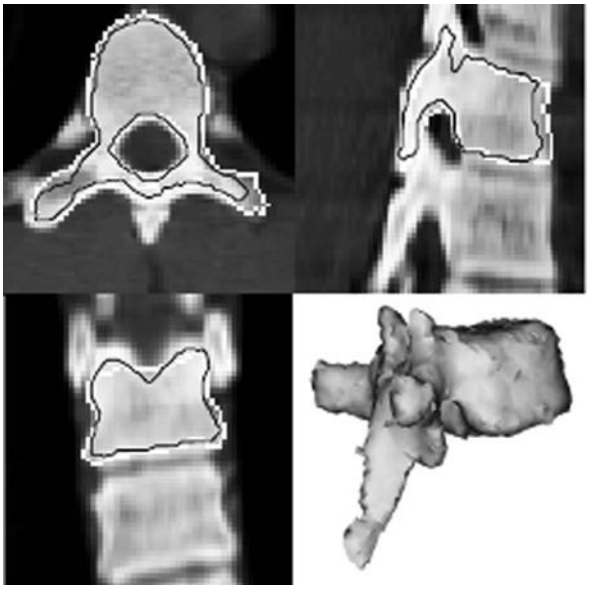
Deformable Models. Figure 1 Tracking a face with the active appearance model. Image from [10].

While building 3D models might be difficult, 3D models naturally offer a better separation of object specific parameters from parameters such as pose and illumination that originate in the specific imaging parameters. For 2D models these parameters are often extremely difficult to separate. For instance, with a 2D model, 3D pose changes can only be modeled by shape parameters. Similarly, 3D illumination effects are modeled by texture variations.

Applications

Deformable Models have found a wide range of applications in many fields of computer science. For biometrics, the most obvious and well-researched applications are certainly face tracking ([10], Fig. 1) and face recognition, [11]. For face recognition, the fact is exploited that an individual face is represented by its shape and texture coefficients. Faces can be compared for recognition or verification by comparing these coefficients.

Another important area in which Deformable Models have found application is in medical image analysis, most importantly medical image segmentation, ([12], Fig. 2).



Deformable Models. Figure 2 3D level set segmentation with shape prior of a vertebrae. Image from [12]. (© 2000 IEEE).

Recent Developments

While the level-set methods allow for greater topological flexibility, the Active Appearance Model and the 3DMM in turn provide an internal energy term representing prior knowledge about the object class. It is natural to combine the advantages of all these methods by using the level-set representation and its resulting external energy term together with the internal energy term incorporating statistical prior knowledge. In [12], Leventon et al. propose such a method that relies on the level-set representation of snakes introduced by Caselles et al. [4]. The internal energy is given by statistical prior knowledge computed directly from a set of level-set functions (distance functions) representing the curves using a standard PCA approach.

Summary

Deformable models provide a versatile and flexible framework for representing a certain class of objects by specifying a model of the object together with its variations. The variations are obtained by deforming the model in accordance to problem specific constraints the deformation has to fulfill. These

constraints represent the prior knowledge about the object and can range from simple smoothness assumption on the deformed object to the requirement that the resulting object still belongs to the same object class. The analysis of novel objects is done by fitting the deformable model to characteristics of a new object. The fitting ranges from simple approaches of matching the object's boundary in an image, to optimally matching the object's full texture. Because of their flexibility, deformable models are used for many applications in biometrics and the related fields of computer vision and medical image analysis. Among others, the most successful use of these models are in automatic segmentation and image analysis and synthesis [3].

Related Entries

- ▶ Active (Contour, Shape, Appearance) Models
- ▶ Face Alignment
- ▶ Face Recognition, Overview
- ▶ Image Pattern Recognition

References

1. Vetter, T., Poggio, T.: Linear object classes and image synthesis from a single example image. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 733–742 (1997)
2. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. Int. J. Comput. Vis. **1**(4), 321–331 (1988)
3. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3d faces. In: SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques, pp. 187–194. ACM Press (1999). DOI http://doi.acm.org/10.1145/311535.311556
4. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. Int. J. Comput. Vis. **22**(1), 61–79 (1997)
5. Cootes, T., Taylor, C.: Active shape models-'smart snakes'. In: Proceedings British Machine Vision Conference, pp. 266–275 (1992)
6. Bishop, C.: Pattern recognition and machine learning. Springer (2006)
7. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. **10**(2), 266–277 (2001)
8. Mumford, D., Shah, J., for Intelligent Control Systems (US, C.: Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems. Center for Intelligent Control Systems (1988)
9. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 681–685 (2001)
10. Matthews, I., Baker, S.: Active Appearance Models Revisited. Int. J. Comput. Vis. **60**(2), 135–164 (2004)
11. Blanz, V., Vetter, T.: Face recognition based on fitting a 3 D morphable model. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1063–1074 (2003)
12. Leventon, M.E., Grimson, W.E.L., Faugeras, O.: Statistical shape influence in geodesic active contours. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn. **1**, 1316 (2000). DOI http://doi.ieeecomputersociety.org/10.1109/CVPR.2000.855835

Deformation

Since expressions are common in faces, robust face tracking methods should be able to perform well inspite of large facial expressions. Also, many face trackers are able to estimate the expression.

- ▶ Face Tracking

Delta

The point on a ridge at or nearest to the point of divergence of two type lines, and located at or directly in front of the point of divergence.

- ▶ Fingerprint Templates

Demisyllables

A demisyllable brackets exactly one syllable-to-syllable transition. Demisyllable boundaries are positioned during the stationary portion of the vowel where there is minimal coarticulation. For general American English, a minimum of 2,500 demisyllables are necessary.

- ▶ Voice Sample Synthesis

Dental Biometrics

Hong Chen, Anil K. Jain
Michigan State University, East Lansing, MI, USA

Synonyms

Dental identification; Forensic Identification Based on Dental Radiographs; Tooth Biometrics

Definition

Dental biometrics uses information about dental structures to automatically identify human remains. The methodology is mainly applied to the identification of victims of massive disasters. The process of dental identification consists in measuring dental features, labeling individual teeth with tooth indices, and the matching of dental features. Dental radiographs are the major source for obtaining dental features. Commonly used dental features are based on tooth morphology (shape) and appearance (gray level).

Motivation

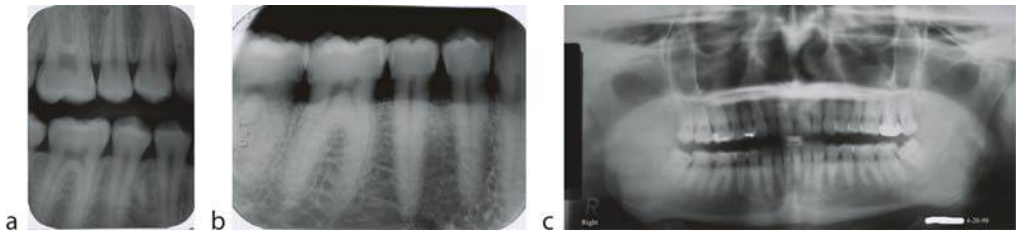
The significance of automatic dental identification became evident after recent disasters, such as the 9/11 terrorist attack in the United States in 2001 and the Asian tsunami in 2004. The victims' bodies were seriously damaged and decomposed due to fire, water, and other environmental factors. As a result, in many cases, common biometric traits, e.g., fingerprints and faces, were not available. Therefore dental features may be the only clue for identification. After the 9/11 attack, about 20% of the 973 victims identified in the first year were identified using dental biometrics [1]. About 75% of the 2004 Asian tsunami victims in Thailand were identified using dental records [2]. Table 1 gives a comparison between dental biometrics and other victim identification approaches, i.e., ▶ circumstantial identification, external identification, internal identification, and ▶ genetic identification [3]. The number of victims to be identified based on dental biometrics is often very large in disaster scenarios, but the traditional manual identification based on forensic odontology is time consuming. For example, the number of Asian tsunami victims identified during the first 9 months was only 2,200 (out of an estimated total of 190,000 victims) [2]. The low efficiency of manual methods for dental identification makes it imperative to develop automatic methods for matching dental records [4, 5].

Dental Information

Dental information includes the number of teeth, tooth orientation, and shape of dental restorations, etc. This information is recorded in dental codes, which are symbolic strings, describing types and positions of dental restorations, presence or absence of each tooth, and number of cusps in teeth, etc. Adams concluded from his analysis [7] that when adequate antemortem (AM) dental codes are available for comparison with postmortem (PM) dental codes, the inherent variability of the human dentition could accurately establish identity with a high degree of confidence. Dental codes are entered by forensic odontologists after carefully reading dental radiographs. Dental radiographs, also called dental X-rays, are X-ray images of dentition. Compared to dental codes, dental radiographs contain richer information for identification, and, therefore, are the most commonly used source of information for dental identification.

Dental Biometrics. Table 1 A comparison of evidence types used in victim identification [6]

Identification approach	Circumstantial	Physical			
		External	Internal	Dental	Genetic
Accuracy	Med.	High	Low	High	High
Time for identification	Short	Short	Long	Short	Long
Antemortem record availability	High	Med.	Low	Med.	High
Robustness to decomposition	Med.	Low	Low	High	Med.
Instrument requirement	Low	Med.	High	Med.	High



Dental Biometrics. Figure 1 Three types of dental radiographs. (**a**) A bitewing radiograph; (**b**) a periapical radiograph; (**c**) a panoramic radiograph.

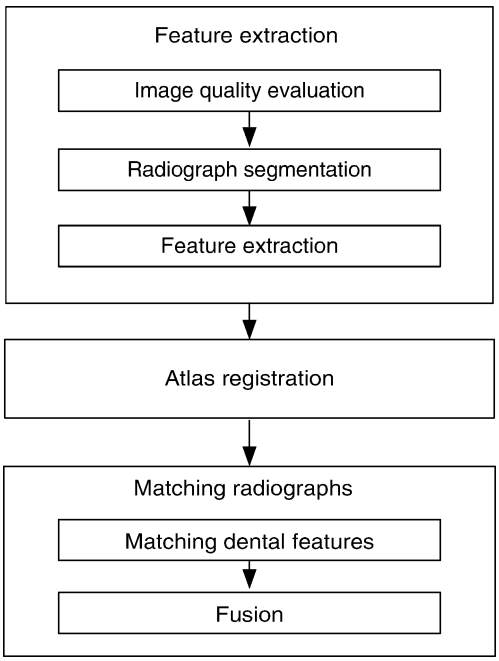
There are three common types of dental radiographs: *periapical*, *bitewing*, and *panoramic*. Periapical X-rays (Fig. 1a) show the entire tooth, including the crown, root, and the bone surrounding the root. Bitewing X-rays (Fig. 1b) show both upper and lower rows of teeth. Panoramic X-rays (Fig. 1c) give a broad overview of the entire dentition (the development of teeth and their arrangement in the mouth), providing information not only about the teeth, but also about upper and lower jawbones, sinuses, and other tissues in head and neck. Digital radiographs will be used for victim identification in future due to their advantages in speed, storage, and image quality.

Antemortem Dental Records

Forensic identification of humans based on dental information requires the availability of antemortem dental records. The discovery and collection of antemortem records is ordinarily the responsibility of investigative agencies. Antemortem dental radiographs are usually available from dental clinics, oral surgeons, orthodontists, hospitals, military service, insurance carriers, and the FBI National Crime Information Center (NCIC).

Automatic Dental Identification System

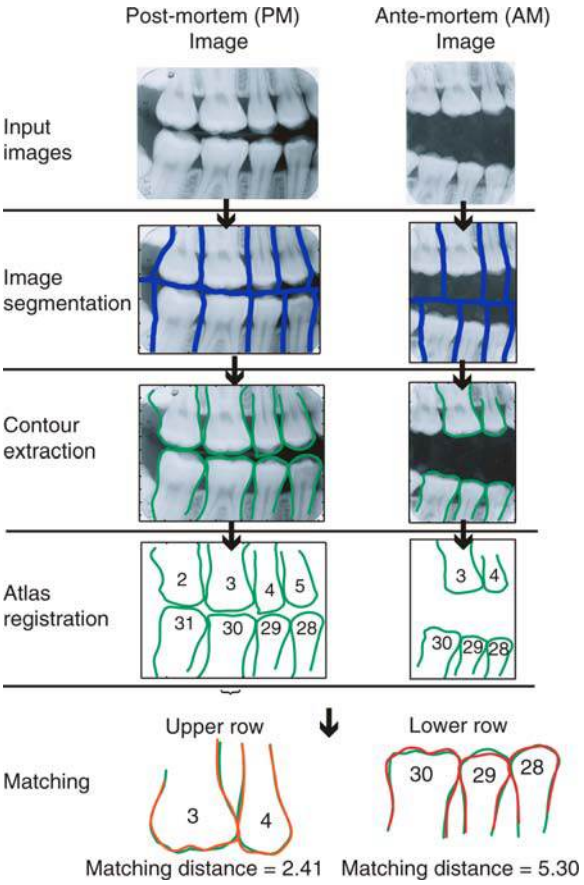
Figure 2 shows the system diagram of an automatic dental identification system [6]. The system consists of three steps: extraction of features, registration of dentition to a dental atlas, and matching dental features. The first step involves image quality evaluation, segmentation of radiographs, and extraction of dental information from each tooth. The second step labels teeth in dental radiographs so that only the corresponding teeth are matched in the matching stage. Dental features extracted from PM and AM images are matched in the third step. For many victims that need to be identified, several AM and PM dental radiographs are available. Therefore, there can be more than one pair of corresponding teeth. In such cases, the matching step also fuses matching scores for all pairs of corresponding teeth to generate an overall matching score between the AM and PM images. Figure 3 shows the process of matching a pair of AM and PM dental radiographs.



Dental Biometrics. Figure 2 Block diagram of automatic dental identification system.

Feature Extraction

The first step in processing dental radiographs is to segment dental radiographs into regions, each containing only one tooth. Segmentation of dental radiographs uses Fast Marching algorithm [8] or integral projection [9, 10]. Figure 4 shows examples of successful radiograph segmentation by the Fast Marching algorithm. Due to the degradation of X-ray films over time as well as image capture in field environments, AM and PM radiographs are often of poor quality, leading to segmentation errors. To prevent propagation of errors in segmentation, an image quality evaluation module is introduced [6]. If the estimated image quality is poor, an alert is triggered to get human experts involved during segmentation.

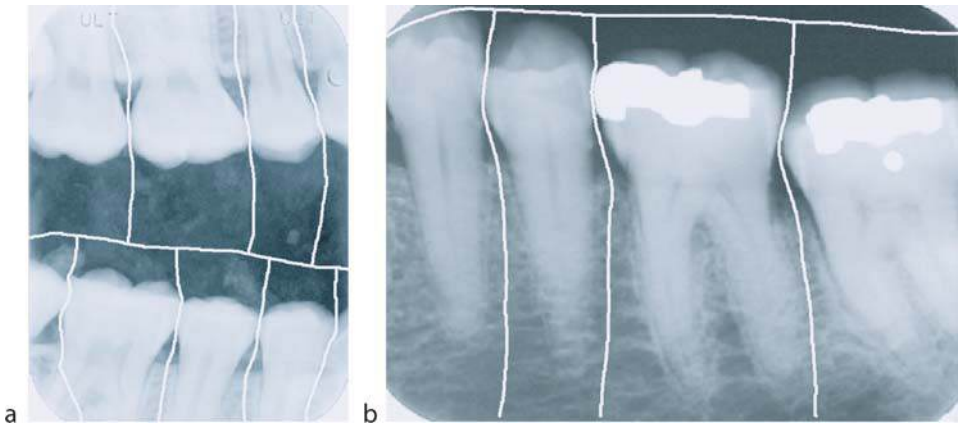


Dental Biometrics. Figure 3 Matching a pair of PM and AM images.

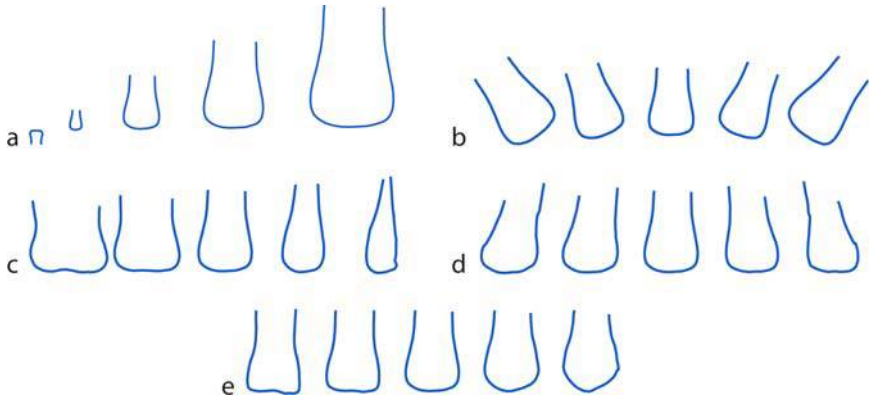
Dental features are extracted from each tooth. The most commonly used features are the contours of teeth and the contours of dental work. Active shape models are used to extract eigen-shapes from aligned training tooth contours [6]. Figure 5 shows the five most principal deformation modes of teeth, which, respectively, represent scaling, rotation, variations in tooth width, variations in tooth orientation, and variations in shapes of tooth root and crown. Figure 6 shows some extracted contours. Anisotropic diffusion is used to enhance radiograph images and segment regions of dental work (including crowns, fillings, and root canal treatment, etc.) [9]. Thresholding is used to extract the boundaries of dental work (Fig. 7).

Atlas Registration

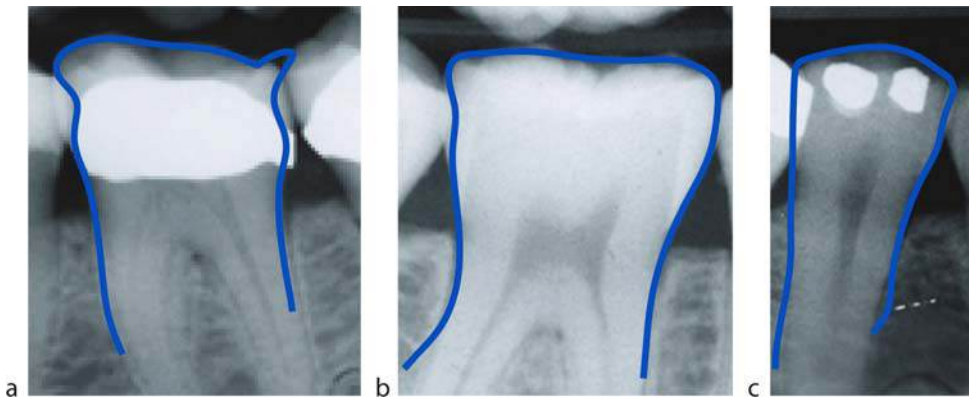
The second step is to register individual teeth segmented in radiographs to a ▶ human dental atlas (Fig. 8). This allows for labeling the teeth with tooth indices. Nomir and Abdel-Mottaleb [11] proposed to form a symbolic string by concatenating classification results of the teeth and match the string against known patterns of



Dental Biometrics. Figure 4 Some examples of correct segmentation.



Dental Biometrics. Figure 5 First five modes of the shape model of teeth. The middle shape in each image is the mean shape, while the other four shapes are, from left to right, mean shape plus four eigenvectors multiplied by −2, −1, 1, and 2 times the square root of the corresponding eigenvalues.

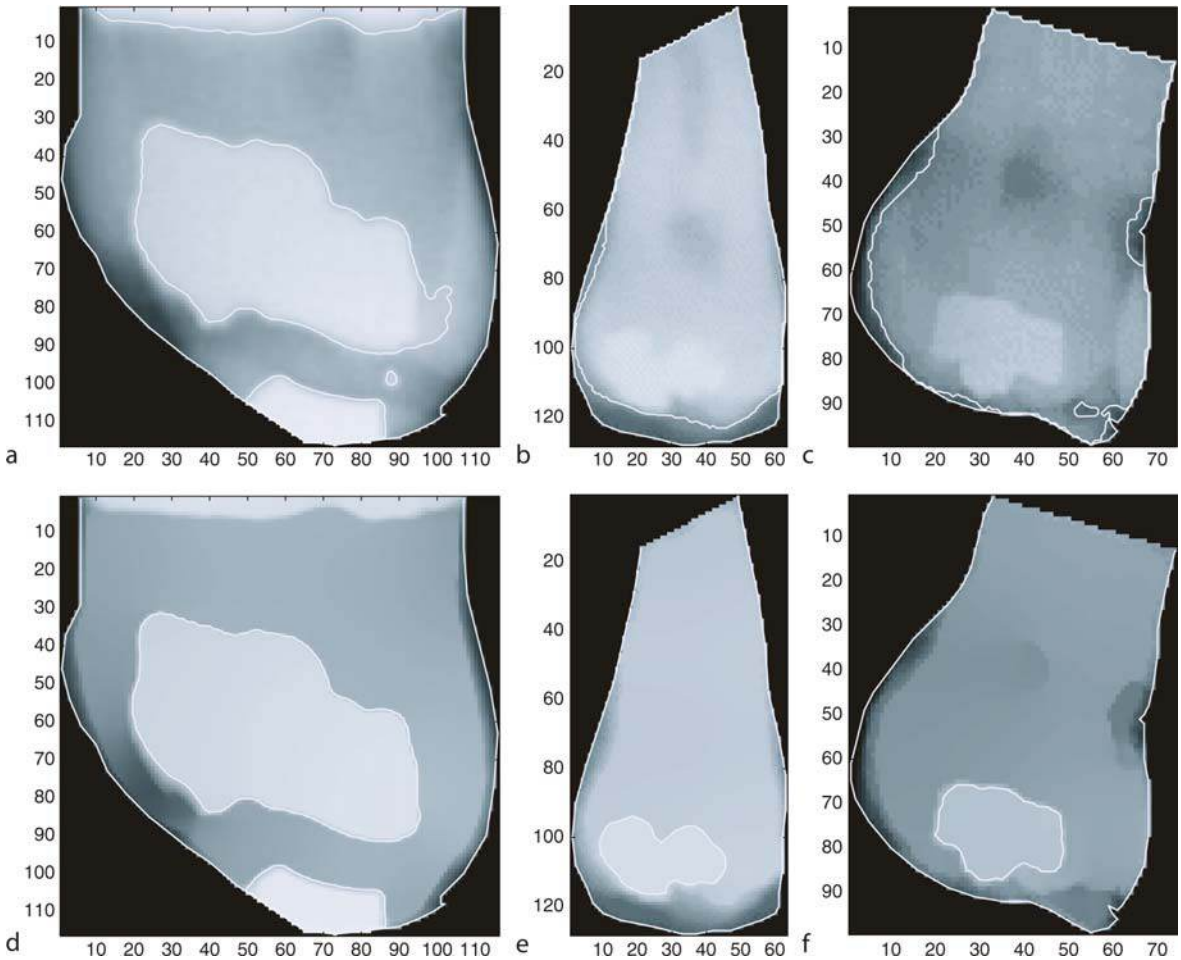


Dental Biometrics. Figure 6 Tooth shapes extracted using Active Shape Models.

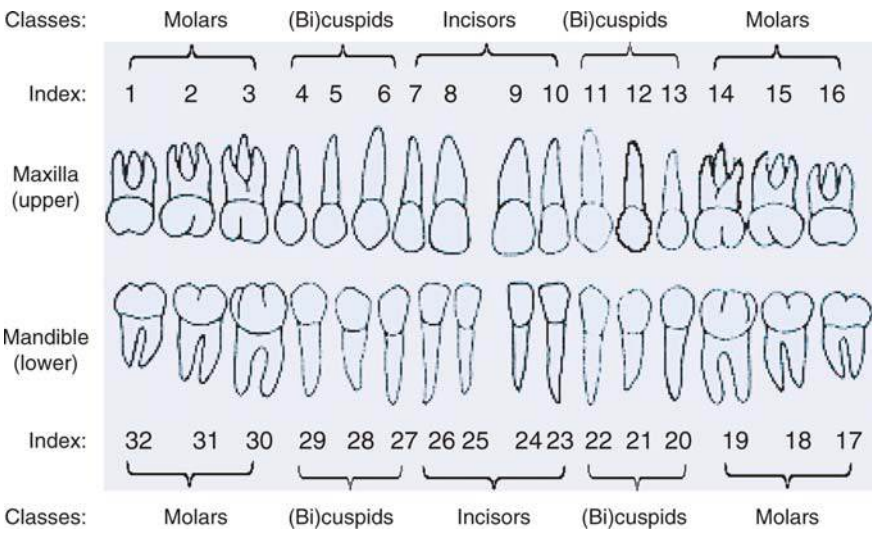
tooth arrangement to find out tooth indices. To handle classification errors and missing teeth, Chen and Jain [12] proposed a hybrid model composed of Support Vector Machines (SVMs) and a Hidden Markov Model (HMM) (Fig. 9). The HMM serves as an underlying representation of the dental atlas, with HMM states representing teeth and distances between neighboring teeth. The SVMs classify the teeth into three classes based on their contours. Tooth indices, as well as Missing teeth, can be detected by registering the observed tooth shapes and the distances between adjacent teeth to the hybrid model. Furthermore, instead of simply assigning a class label to each tooth, the hybrid model assigns a probability of correct detection to possible indices of each tooth. The tooth indices with the highest probabilities are used in the matching stage. Figure 10 shows some examples of tooth index estimation.

Matching

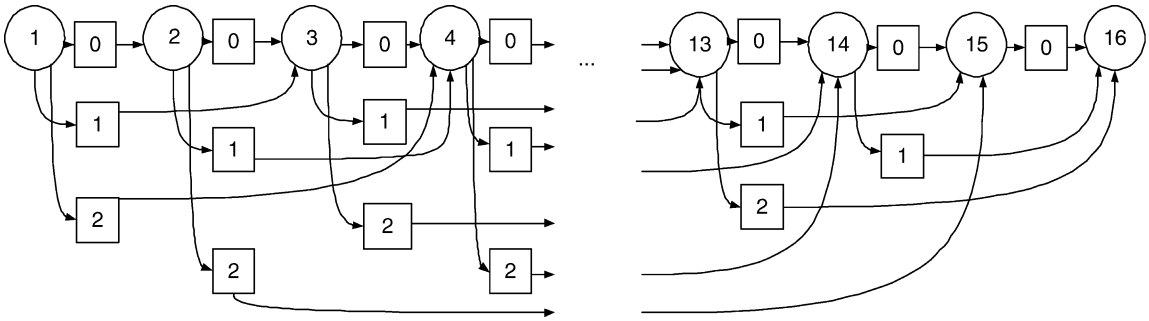
For matching the corresponding teeth from PM and AM radiographs, the tooth contours are registered using scaling and a rigid transformation, and corresponding contour points are located for calculating the distance between the contours [9]. If dental work is present in both the teeth, the regions of dental work are also matched to calculate the distance between dental work. The matching distance between tooth contours and between dental work contours are fused to generate the distance between the two teeth [9]. Given the matching distances between individual pairs of teeth, the matching distance between two dental radiographs is computed as the average distance of all the corresponding teeth in them. The distance between the query and a record in the database is calculated based on the distance between all the available dental radiographs in the



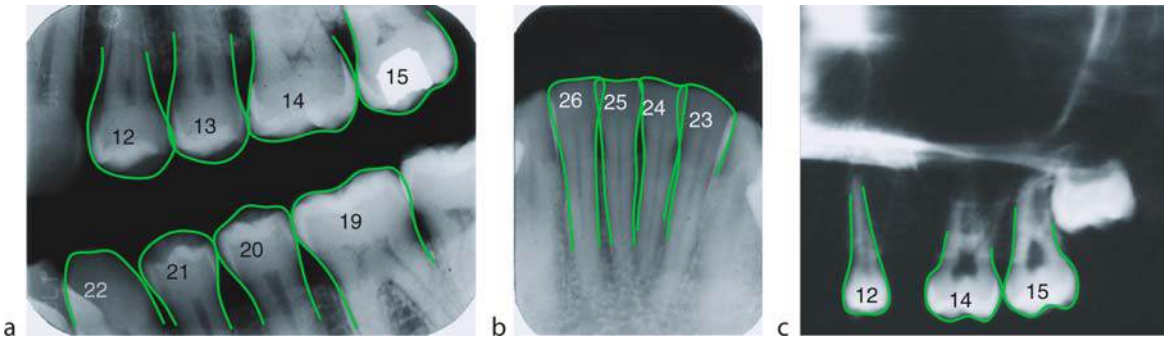
Dental Biometrics. **Figure 7** Extracted dental work contours with and without image enhancement. (**a**), (**b**) and (**c**) Without enhancement. (**d**), (**e**) and (**f**) After enhancement.



Dental Biometrics. **Figure 8** Dental Atlas of a complete set of adult teeth containing indices and classification labels of the teeth.



Dental Biometrics. Figure 9 SVM/HMM model for the upper row of 16 teeth. The circles represent teeth, and the number inside each circle is the tooth index. The squares represent missing teeth, and the number inside each square is the number of missing teeth.



Dental Biometrics. Figure 10 Examples of successful registration of the dental atlas to (**a**) a bitewing image, (**b**) a periapical image, and (**c**) an image with a missing tooth. In (**c**), teeth numbered as 12, 14, and 15 are correctly registered. The missing tooth (number 13) is detected.

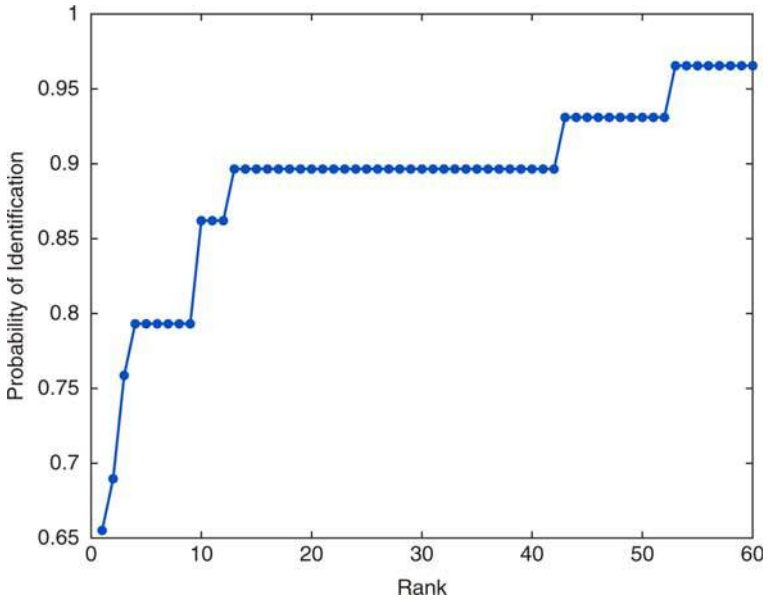
query and the database record. The distance between the query and all the database record is used to find the closest match for a given query, and according to the rank, an ordered list of candidate identities is generated.

Performance Evaluation

To the authors' knowledge, no public domain databases are available to evaluate the performance of automatic dental matching. Hong and Jain [6] conducted experiments on a dental X-ray database consisting of 29 PM subjects and 133 AM subjects [6]. There are 360 PM tooth sequences (810 teeth) and 1,064 AM tooth sequences (3,271 teeth) in total. Figure 11 shows the Cumulative Match Characteristics (CMC) curve based on this database. The accuracy for top-1 retrieval is 19/29 (66%), the accuracy for top-4 retrievals is 23/29 (79%), and the accuracy for top-13 retrievals is 26/29 (90%). Most of the errors in retrieval are attributed to change in the appearance of dentition due to loss of teeth, matching different types of teeth, and poor quality of AM and PM images.

Other Approaches

Other matching approaches have been attempted for human identification based on dental radiographs. Hofer and Marana [13] used edit distance to match dental codes extracted from dental works in radiographs. Nikaido et al. [14] proposed to use image registration approach. Nomir and Abdel-Mottaleb [15] compared Fourier descriptors of tooth contours and other features based on gray level information of teeth in images.



Dental Biometrics. **Figure 11** Cumulative matching characteristics (CMC) curves for subject retrieval in a database of 133 subjects [6].

Summary

Massive human disasters make it imperative to research automatic dental identification methods to identify anonymous human remains. Dental radiographs contain valuable clues that often is the only source of information to identify victims. An overview of automatic identification methods based on dental radiographs was given in this entry. The accuracy and efficiency of current approaches need to be further improved. A large database of AM and PM radiographs needs to be collected and made available to researchers to evaluate performance of the automatic systems under development.

Related Entries

- ▶ Feature Extraction
- ▶ Forensic Applications, Overview
- ▶ Hidden Markov Models
- ▶ Support Vector Machine

References

1. O'Shaughnessy, P.: More than half of victims IDd. New York Daily News (11 Sep. 2002)
2. Dental records beat DNA in tsunami IDs. New Scientist **2516,** 12 (2005). Http://www.newscientist.com/article.ns?id=mg18725163.900
3. Disaster victim identification. Http://www.interpol.int/Public/DisasterVictim/Guide
4. Fahmy, G., Nassar, D., Haj-Said, E., Chen, H., Nomir, O., Zhou, J., Howell, R., Ammar, H.H., Abdel-Mottaleb, M., Jain, A.K.: Towards an automated dental identification system (ADIS). In: Proceedings of ICBA (International Conference on Biometric Authentication), vol. LNCS 3072, pp. 789–796. Hong Kong (2004)
5. Jain, A., Chen, H., Minut, S.: Dental biometrics: human identification using dental radiographs. In: 1Proceedings of Fourth International Conference on AVBPA (Audio- and Video-based Biometric Person Authentication), pp. 429–437. Guildford, U.K. (2003)
6. Chen, H., Jain, A.K.: 1Handbook of Biometrics, chap. on Automatic Forensic Dental Identification. Springer, Berlin (2007)
7. Adams, B.: The diversity of adult dental patterns in the United States and the implications for personal identification. J. Forens. Sci. **48**(3), 497–503 (2003)
8. Chen, H.: Automatic forensic identification based on dental radiographs. Ph.D. thesis, Michigan State University (2007)
9. Chen, H., Jain, A.: Dental biometrics: Alignment and matching of dental radiographs. IEEE Trans. Pattern Analy. Mach. Intell **27**(8), 1319–1326 (2005)
10. Nomir, O., Abdel-Mottaleb, M.: A system for human identification from X-ray dental radiographs. Pattern Recogn. **38**(8), 1295–1305 (2005)
11. Mahoor, M.H., Abdel-Mottaleb, M.: Classification and numbering of teeth in dental bitewing images. Pattern Recogn **38**(4), 577–586 (2005)
12. Jain, A., Chen, H.: Registration of dental atlas to radiographs for human identification. In: Proceedings of SPIE Conference on Biometric Technology for Human Identification II, vol. 5779, pp. 292–298. Orlando, Florida (2005)

13. Hofer, M., Marana, A.N.: Dental biometrics: Human identification based on dental work information. In: Proceedings of Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI), pp. 281–286 (2007)
14. Nikaido, A., Ito, K., Aoki, T., Kosuge, E., Kawamata, R.: A dental radiograph registration algorithm using phase-based image matching for human identification. In: Proceedings of International Symposium on Intelligent Signal Processing and Communications, pp. 375–378 (2006)
15. Nomir, O., Abdel-Mottaleb, M.: Human identification from dental X-ray images based on the shape and appearance of the teeth. IEEE Trans. Inform. Forens. Sec **2**(2), 188–197 (2007)

Dental Identification

▶ Dental Biometrics

Deployment

The engineering of technological ideas and devices to become useful by-product in real-industry environments for operators and users.

▶ Biometric Applications, Overview

Depth of Field (DOF)

Depth of Field (DOF) is the distance in front of and beyond the subject that appears to be in focus. It depends on the camera lens.

▶ Face Device
▶ Iris on the Move™

Dermis

Dermis is the inner layer of the skin. It varies in thickness according to the location of skin. It is 0.3 nm on eyelid while 3.0 on back. It is composed of three types of tissues which are present throughout the layers named as: (1) collagen, (2) elastic tissue, and (3) reticular fibers. The two layers of the dermis are the papillary and reticular layers. The upper, papillary layer, which is the outer layer, contains a thin arrangement of collagen fibers. The lower, reticular layer is thicker and made of thick collagen fibers that are arranged parallel to the surface of the skin.

▶ Anatomy of Hand

DET Curves

Detection Error Tradeoff curves are ROC (receiver operating characteristic) type curves showing the range of operating points of systems performing detection tasks as a threshold is varied to alter the miss and false alarm rates and plotted using a normal deviate scale for each axis. DET curves have the property that if the underlying score distributions for the two types of trials are normal, the curve becomes a straight line. They have been widely used to present the performance characteristics of speaker recognition systems.

▶ Speaker Databases and Evaluation

Detector – Extractor

Traditionally, the term detector has been used to refer to the tool that extracts the features from the image, e.g., a corner, blob, or edge detector. However, this only makes sense if it is a priori clears what the corners, blobs, or edges in the image are, so one can speak of "false detections" or "missed detections." This only holds in the usage scenario where features are semantically meaningful, otherwise extractor would probably be more appropriate. The term detector is however more widely used.

▶ Local Image Features

Diffraction Limit

In optics, a fundamental limit on the resolution that can be obtained with an imaging system. The limit depends on the wavelength of the light used, the size of the aperture of the optical system and the distance between the optical system and object being imaged.

► Iris on the Move

Diffuse Reflection

Diffuse reflection is the reflection of light from a rough surface. When a bunch of parallel light reach the rough surface, the reflected light will be spread in all directions. It is the complement to specular reflection.

► Skin Spectroscopy

Digital Watermarking

► Iris Digital Watermarking

Digitizer

► Digitizing Tablet

Digitizing Tablet

Sonia Garcia-Salicetti, Nesma Houmani
TELECOM SudParis, Evry, France

Synonyms

Digitizer; Graphic tablet; Touch tablet; Tablet

Definition

A digitizing tablet is a sensitive input device that converts a hand-drawn trajectory into a digital on-line form, which is a sequence. This hand-drawn trajectory may be a signature input, handwriting, or hand-drawn graphics. A digitizing tablet usually consists of an electronic tablet and a pen or a stylus. When the electronic tablet communicates with the pen, it is said to be "Active" and in this case the pen contains an electronic circuit; otherwise the tablet is said to be "Passive." In some cases, the digitizer only consists of an electronic (Active) pen, used on either standard or special paper. Active ► digitizing tablets sample the pen trajectory at regular time intervals (around 10 ms), generating a time stamp and associated time functions; passive digitizing tablets require dedicated acquisition software to retrieve the sequence of time stamps and associated time functions. When digitizing human pen input, the resulting output may have different forms according to the type of digitizer used. Active digitizing tablets capture a sequence of time functions, including pen position, pen pressure, and pen inclination, while passive digitizing tablets, with acquisition software, only allow a time stamp and the position of the stylus on the tablet to be captured. In the special case of electronic pens, the digitizer may capture some of the previously mentioned time functions plus some others, such as pen acceleration.

Introduction

Digitizing tablets available nowadays can be based on electromagnetic technology (Active digitizing tablets), touch screen technology (Passive digitizing tablets), hybrid technology combining both (respectively called "Active mode" and "Passive mode"), or, finally, in the case of digitizers consisting only of an electronic pen (no tablet, just an Active Pen and a sheet of standard or special paper) on a variety of principles (mechanical, optical). Each kind of digitizing tablet has been discussed here.

Active Digitizing Tablet: Electromagnetic Technology

This is the technology of choice in biometric applications related to online signature verification and writer

authentication by online handwriting, because it is the technology with the highest resolution and accuracy at the acquisition step.

In this case, the digitizing tablet is based on ► electromagnetic resonance technology [1] and it contains electronics external to the touched surface. An active digitizing tablet consists of a sensitive acquisition surface incorporating a large number of overlapping loop coils arranged horizontally and vertically, and a special pen containing an inductor-capacitor circuit (LC circuit). Electromagnetic resonance then allows information exchange between the tablet and the pen in the following way: the tablet generates a magnetic field, which is received by the LC circuit in the pen, and then the pen's resonant circuit makes use of this energy to return an electromagnetic signal to the tablet [1]. Here, the horizontal and vertical wires of the tablet operate as both transmitters and receivers of magnetic signals. This exchange of information allows the tablet to detect the position, pressure, and inclination of the pen.

Using electromagnetic resonance technology in particular allows the pen position to be sensed without the pen even having to touch the tablet (ability to hover) and also the user's hand may rest on the flat acquisition surface without affecting the acquisition process (the capture of the time functions of position, pressure, and inclination of the pen).

Active digitizing tablets are of two types: one in which the tablet powers the pen, thus avoiding the use of batteries in the pen (as in the case of well-known digitizing tablets on the market, for example, Wacom digitizers [1]) and the other in which the pen requires batteries (as in the case of other vendors, such as AceCad or Aiptek).

The ► sampling frequency of Active digitizing tablets may be tuned for acquisition. For example, in Wacom Intuos2 A6 USB tablet, the sampling frequency of the hand-drawn signal can reach 200 Hz and is frequently set around 100 Hz.

Passive Digitizing Tablets: Touch-Screen Technology

In this case, all the electronics are inside the digitizing tablet, based on touch-screen technology, and are activated by a non-sensitive stylus. Passive digitizing tablets are integrated into other multifunction devices like PCs (such as a Tablet PC), or handheld devices, such as Personal Digital Assistants (PDAs) or Smartphones. Specific acquisition software is required in these digitizers in order to retrieve the sequence of time stamps and associated time functions corresponding to the hand-drawn trajectory on the Touch Screen. Passive digitizing tablets allow fewer time functions (only a time stamp and the associated pen position on the Touch Screen), than Active digitizing tablets to be captured. Also, spatial resolution is variable in these digitizing tablets and is less precise than that obtained in Active digitizing tablets.

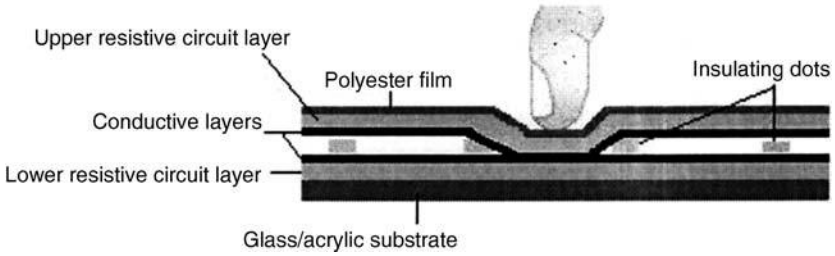
There exist several types of touch-screen technology (resistive, capacitive, infrared, Surface Acoustic Wave, and Near Field Imaging) but the two mostly used for online signature capture and more generally for online handwriting recognition, that is, resistive and Pen-Touch Capacitive technology, are discussed here.

Resistive Passive Digitizing Tablets

The most widely used Passive digitizing tablets today are based on resistive technology [2]. This technology can be summarized by the fact that a touch on the screen generates a tension (a voltage) at a localized point that is at the position of touch [2]. In a Resistive Touch Sensor, there are two upper thin metallic conductive layers separated by a thin space (often filled with tiny spacer dots) in between two resistive layers (Fig. 1). The Resistive Sensor is mounted above a ► liquid crystal display (LCD) to sense pressure. Pressure from using either a stylus or a finger bends the upper conductive layer, which is flexible, producing an electrical contact between the two conductive layers received by the LCD. A controller locates the pen by measuring the tension between the point of contact and the edges or corners of the touch screen. Resistive Sensors cannot distinguish pens from fingers and do not have the hover ability of Active digitizing tablets (cannot detect proximity of the pen or finger without actual pressure).

Pen-Touch Capacitive Passive Digitizing Tablets

This technology can be summarized by the fact that a touch with a tethered stylus (special stylus with a conductive tip) reduces electrical charges on the



Digitizing Tablet. Figure 1 Principle of resistive touch-screen technology.

screen [2]. A layer that stores electrical charges is placed on the glass panel of the monitor. When the stylus touches the screen, some of these charges are transferred to the stylus and thus, the charge on the capacitive layer decreases. This decrease is measured in circuits located at each corner of the glass panel, and the information relayed by the computer to the touch screen driver software to determine the coordinates of the touch.

This technology has three modes: the screen can be set to respond to finger input only, pen input only, or both. The pen stylus is used in particular for signature capture and offers online handwriting recognition facilities. One popular touch screen using this technology is ClearTek II produced by 3M Touch Systems [3], sometimes integrated in monitors (LCDSA121-PEN-S-OF [4]); other examples of this technology are Apple PDAs, such as iPhone and iPod (in particular, Songtak Technology CoLtd [5] designed a special stylus with a conductive tip for iPhone and iPod).

Hybrid Digitizing Tablets: Active Mode and Passive Mode

Some digitizing tablets are hybrid as they can operate in both Active and Passive modes. There are cases in which such digitizing tablets combine capacitive and electromagnetic technologies, such as the ClearPad and Spiral sensors of Synaptics [6] (one of the leaders in capacitive touch-sensing technology) and others in which resistive and electromagnetic technologies are combined, such as the Tablet-PC Sahara slate PC i440D [7]. These hybrid digitizing tablets can distinguish pen input from finger input since, when there is a magnetic field, what produces the touch is an Active Pen. Microsoft's Tablet PCs belong to this category of hybrid digitizing tablets.

Digitizing Tablets with Active Pen Only

In this case, the digitizing tablet is completely mobile, since it consists of a sensitive pen, which is like a normal ballpoint pen in shape as well as grip. Such devices even use ink for writing but, by means of different principles, a sequence of data from the hand-drawn signal is stored in the memory of the device.

There are different categories of Active Pens, according to the underlying principles allowing the acquisition of a sequence of data from a hand-drawn signal. The first category is based on a ▶ piezoelectric element that transforms a mechanical force into an electrical signal (voltage); this category is represented for instance by the "Marking Device" [8]. Another category is based on ▶ strain gauges that transform their deformation (strain) into a change in an electrical signal, usually a resistive circuit; this category is represented by the SmartPen [9]. The third category is based on optical sensors [10–15].

In the Electronic pen [8], the Marking Device includes a pressure sensor and two acceleration sensors. The pressure sensor is coupled to the tip of the pen. The acceleration sensors adjacent to the pen tip sense acceleration of the tip in two directions. These sensors are based on piezoelectric transducers; for example, the pressure sensor has at least one pair of electrodes coupled to a piezoelectric element that is compressed as a result of the pressure exerted on the tip. The piezoelectric transducer then conveys the resulting compression into an electrical signal. From the sequence of pressure values, pen acceleration and

temporal information related to the sampling of the acceleration sensors, the Marking Device extracts other features, such as speed, position, and angular information about the acceleration vector [8].

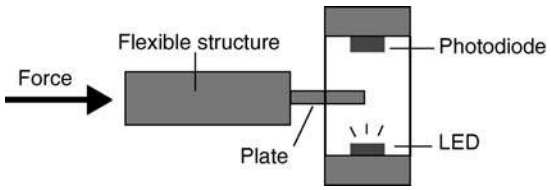
In the second category, a mechanical force is transformed into an electrical signal, by the use of strain gauges. An example of this technology is the SmartPen [9], made by the LCI Computer Group. The pen contains micro-electromechanical sensors, one sensing both force (pressure) and acceleration, the other sensing tilt, and a radio transmitter to send the information to a computer. The LCI-SmartPen serves to authenticate the signer of a document through the dynamics of the signature. While the signer writes, micro-sensors measure the inclination and pressure of the pen and the acceleration of the pen tip in three directions by means of a ring-shaped deformable aluminium structure provided with strain gauges. The information is processed and then transmitted to a computer. The position of the pen and the pen tip velocity are calculated through measurements obtained from the force/acceleration sensor and the tilt sensor.

In the third category, optical principles are used to acquire a data sequence from the hand-drawn signal.

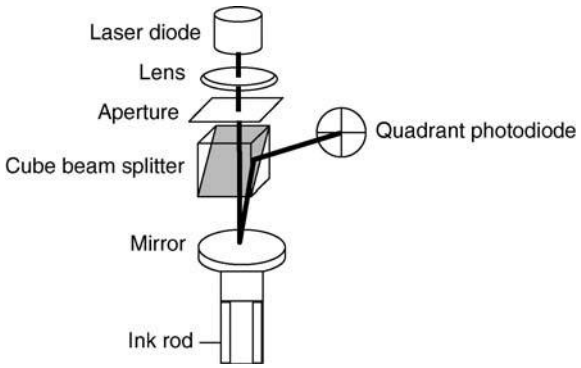
The first type of Active Pen in this category relies on a digital camera integrated in the pen and special paper; this is the Anoto Active Pen (of Anoto AB Company) [10, 11]. When writing with the Active Pen on Anoto paper, which contains a special pattern of numerous black dots, digital snapshots are taken automatically by an integrated digital camera (more than 50 pictures per second), and the dots of the written pattern are illuminated by infrared light, making them visible to the digital camera. In this way, a sequence of information (timing, coordinates, etc.) is captured from the hand-drawn signal. With the very same principle, it is found that the Sony Ericsson Chatpen [11] and the Logitech io pen [12] also require the Anoto special paper.

The second type of Active Pen in this category relies on ▶ diodes; an example is the V-Pen using an Optical Translation Measurement sensor (OTM sensor), including a laser diode, detectors, and optics integrated into a small transistor-style package [13]. The laser diode shines laser beams on to the writing surface and the OTM sensor analyzes how movements of the pen affect the reflected wavelengths [13]. The relative motion in three dimensions is measured based on the Doppler Effect – the change in frequency and wavelength of a wave as perceived by an observer moving relative to the source of the waves. The V-Pen performs signature recognition by combining the dynamic characteristics of the signing action to the signature shape.

Other Active Pens in this category have appeared as research prototypes [14, 15]. The first measures pen pressure (force) in three dimensions by Optical principles [14]. The pressure sensor uses a flexible element in combination with an optical displacement sensor based on ▶ light emitting diodes (LED) and photodiodes. The pressure on the pen tip is applied to a flexible structure that deforms. A plate, placed between the infrared LED and the photodiode, moves together with the deformed structure (Fig. 2). Its movement changes the amount of incoming light on the photodiode and, therefore, the electrical current of the Photodiode (that corresponds to a pressure value). In a new device [15], pen inclination is measured by means of a laser diode used as a light source. The laser beam is reflected by a mirror mounted on the end of the ink rod. The reflected beam is guided to a



Digitizing Tablet. **Figure 2** Principle of the optical force sensor.



Digitizing Tablet. **Figure 3** Principle of the optical force sensor.

Quadrant Photodiode (QPD) via a cube beam splitter (Fig. 3). When writing, the mirror attached to the ink rod is tilted in X and Y directions and the path of the reflected beam changes. Consequently, the optical spot on the photodiode moves, causing variations in the output currents of the QPD cells. The sampling frequency of the output voltages is set at 200 Hz and the resulting sequence of voltages is interpreted as pen inclination information.

Related Entries

- ▶ Acceleration
- ▶ Altitude
- ▶ Azimuth
- ▶ Biometric Sensor and Device, Overview
- ▶ Feature Extraction
- ▶ On-Line Signature Verification
- ▶ Pen Inclination
- ▶ Pen Pressure
- ▶ Position
- ▶ Sensors
- ▶ Signature Features
- ▶ Signature Recognition
- ▶ Timing
- ▶ Velocity

References

1. http://www.wacom.com/productinfo/. Accessed 6 Dec 2007
2. Synaptics Incorporated, New Touch Screens Improve Handheld Human Interface. Synaptics White Paper WP-12, P/N 507–000003, Rev A. Synaptics Incorporated, San Jose California
3. solutions.3m.com/wps/portal/3M/en_US/3MTouchSystems/TS/. Accessed 14 Dec 2007
4. www.touchscreens.com/lcdsa121-pen-s-kr.html. Accessed 14 Dec 2007
5. www.songtak.com.tw/. Accessed 6 Dec 2007
6. www.synaptics.com/onyx/. Accessed 6 Dec 2007
7. www.tabletkiosk.com/products/sahara/i400s_pp.asp. Accessed 20 Dec 2007
8. O'Connor, M., Vannier, D.S.: Electronic pen device. US Patent 6,188,392, 13 Feb 2001
9. Reynaerts, D., Peirs, J., Van Brussel, H.: A mechatronic approach to microsystem design. IEEE/ASME Trans. Mechatronics **3**(1), 24–33 (1998)
10. Skantze, K.: Method and device for secure wireless transmission of information. US Patent 7,278,017, 2 Oct 2007
11. Euchner, J.A., Coffy, J.H., Obrea, A.: System and method for annotating documents. US Patent 7,111,230, 19 Sept 2006
12. www.fruits-it.com/request.php?8. Accessed 20 Dec 2007
13. http://www.otmtech.com. Accessed 6 Dec 2007
14. Clijnen, J., Reynaerts, D., Van Brussel, H.: Design of an optical tri-axial force sensor. In: Bar-Cohen, Y. (ed.) Proceedings of SPIE, vol. 4946, Transducing Materials and Devices, March 2003, pp. 129–136
15. Shimizu, H., Kiyono, S., Motoki, T., Gao, W.: An electrical pen for signature verification using a two-dimensional optical angle sensor. Sensor Actuator **111**, 216–221 (2004)

Dimensionality Reduction

The process of reducing the number of features used by a classifier in order to decrease measurement cost, increase classification accuracy, and mitigate the problems associated with the curse-of-dimensionality. Feature selection or feature extraction techniques is used to deduce an optimal (or, from a practical standpoint, sub-optimal) set of features from a pool of available features. Common examples include sequential forward selection (SFS), sequential backward selection (SBS), sequential forward floating search (SFFS), etc.

- ▶ Fingerprint Sample Synthesis
- ▶ Fusion, Feature-Level

Diode

Diode is an electronic component that allows the passage of current in only one direction. A light emitting diode (LED) is an electronic semiconductor diode that emits a single wavelength of light when electric current passes through it. A photodiode is a semiconductor diode that allows current to flow when it absorbs photons (light).

- ▶ Digitizing Tablet

Diphones

A diphone brackets exactly one phoneme-to-phoneme transition. Diphone boundaries are usually positioned

near the midpoint of the most stationary (non-changing) region of two consecutive phonemes. Theoretically, a phoneme inventory of 50 could give rise to up to 2,500 diphones, but not all diphones exist in a given language. For general American English, a minimum of about 1,500 diphones are necessary.

► Voice Sample Synthesis

Disclosure Check

► Background Checks

Discriminative Classifier

A discriminative classifier is a classification algorithm than learns a borer; one side it labels one class, the other side it labels another. The border is chosen to minimize error rate, or some correlated measure, effectively discriminating between the classes.

► Fusion, Quality-Based

Dissimilarity

► Palmprint Matching

Distortion

Variances in the ridges and features of a fingerprint caused by deposition pressure or movement are distortion.

► Universal Latent Workstation

Distributed Computing

Distributed computing refers to the paradigm of not having a central computing node in a sensor network. Distributed computing uses parallel computations over a multiple processing units, connected by a communications network. Distributed computing alleviates the need for having an extremely powerful central computing node in a network, sacrificing performance to obtain robustness to partial failure of the network both in terms of individual node failures as well as failure on the communication links.

► Surveillance

Distributed Detection

► Fusion, Decision-Level

Distributed Inference Making

► Fusion, Decision-Level

DNA Analysis

► Forensic DNA Evidence

DNA Fingerprinting/DNA Profiling

DNA fingerprinting is a term coined by Sir Alec Jeffreys describing the multi-locus probes results obtained in

1985 (i.e., bar code type output). The analogy with fingerprint should be avoided, and the term DNA profiling suggested by Evett & Buckleton preferred. Indeed, the term profile indicates that this type of analysis does not allow characterizing a person's DNA, but only given parts called markers. An explicit mention of the markers and the technique used should always accompany a given DNA profile.

► Forensic DNA Evidence

DNA Profiling

► Forensic DNA Evidence

DNA Typing

► Forensic DNA Evidence

Dolicocephalic

Dolicocephalic is the head form characterized by an anteroposteriorly long and mediolaterally narrow skull.

► Anatomy of Face

Double Angle Representation

It refers to doubling the angles of the gradients making them fit to represent ridge directions continuously.

► Fingerprint Features

Double Dipping

Double dipping refers to the unethical act of seeking compensation, benefits, or privileges from one or more sources, given only a single legitimate entitlement. In the context of biometrics, double dipping usually occurs when an individual seeks such unauthorized advantage and/or gain by assuming multiple nominal identities.

► Fraud Reduction, Applications

Drive-up

► Iris on the Move™

Duplicate Detection

► Fraud Reduction, Applications

Dynamic Programming Comparison Method

Method of mathematical programming developed by Richard Bellman (Dynamic Programming, Princeton University Press, 1957).

It is useful in obtaining the optimal strategy when the objective function to be maximized is monotonic and recursively defined depending on a variable. In the dynamic programming comparison method for signature recognition, dynamic programming is applied to obtain the best warping function of the location variable of the template against the location variable of the questionable image, using the similarity between the questionable image and the template as the objective function.

► Signature Recognition

Dynamic Time Warping (DTW)

The DTW is a method used in text-dependent Speaker Recognition. In this context, the training or testing data are composed by a sequence of acoustic vectors and the temporal order of the vectors is important. In order to compute likelihood or a distance between two of such sequences, two functions are needed, a frame to frame distance function and a frame mapping function, able to align the individual acoustic frames of both sequences. This time alignment function is mandatory as two occurrences of the same linguistic messages, pronounced or not by the same speaker, present different time characteristics, like the global pronunciation speed. If there is a training template T_r with N_{T_R} frames and a test utterance T_E consisting in a sequence of N_{T_E} frames, the DTW is able to find the time mapping function $w(n)$ between T_R and T_E. In the figure 1, the tying function $w(n)$ is illustrated by the tying of the T_R frame at time x with the T_E frame at time y. Thus, the system can evaluate a distance D() between T_R and T_E, defined by the following formula:

$$D\left(T_R T_E = \frac{1}{N_{T_R}} \sum_{k=1}^{N_{T_R}} d\left(v_n^{T_R}, v_{w(n)}^{T_E}\right)\right),$$

where, $v_n^{T_R}$ is a acoustic vector (cepstral vector) of the training message TR at time n, $v_{w(n)}^{T_R}$ the time-aligned acoustic vector of the test message TE and $d()$ is the frame to frame distance. The distance estimated in Eq. 11 (thanks to DTW algorithm) is used to make the decision of accepting or rejecting the claimed identity.

▶ Signature Matching
▶ Speaker Matching

E

Ear Biometrics

Michał Choraś
Institute of Telecommunications University of Technology and Life Sciences,
Bydgoszcz, Poland

Synonym

Ear Recognition

Introduction

Biometrics identification methods have proved to be very efficient, more natural and easy for users than traditional methods of human identification. Biometrics methods truly identify humans, not keys and cards they posses or passwords they should remember. The future of biometrics leads to systems based on image analysis as the data acquisition is very simple and requires only cameras, scanners or sensors. More importantly, such methods could be ▶ passive, which means that the subject does not have to take active part in the whole process or, in fact, would not even know that the process of identification takes place. There are many possible data sources for human identification systems, but the physiological biometrics has many advantages over methods based on human behavior. The most interesting human anatomical parts for passive, physiological biometrics systems are face and ear.

There are many advantages of using the ear as a source of data for human identification. Firstly, the ear has a very rich structure of characteristic ear parts. The location of these characteristic elements, their direction, angles, size and relation within the ear are distinctive and unique for humans, and therefore, may be used as a modality for human identification [1, 2].

Ear is one of the most stable human anatomical feature. It does not change considerably during human life while face changes more significantly with age than any other part of human body [1, 2]. Face can also change due to cosmetics, facial hair and hair styling. Secondly, human faces change due to emotions and express different states of mind like sadness, happiness, fear or surprise. In contrast, ear features are fixed and unchangeable by emotions. The ear is not symmetrical – the left and right ears are not the same. Due to forensics and medical studies, from the age of 4 ears grow proportionally, which is the problem of scaling in computer vision systems [1].

Furthermore, the ear is a human sensor, therefore it is usually visible to enable good hearing. In the process of acquisition, in contrast to face identification systems, ear images cannot be disturbed by glasses, beard or make-up. However, occlusion by hair or earrings is possible.

It is also important that ear biometrics is highly accepted biometrics by users in possible access control applications and government security such as visa/passport programs. According to users, ear biometrics is less stressful than fingerprinting. Moreover, users admitted that they would feel less comfortable while taking part in face images enrolment (people tend to care how they look on photographs) [3]. Furthermore, in ear biometrics systems there is no need to touch any devices and therefore there are no problems with hygiene.

It is worth mentioning that ear images are more secure than face images, mainly because it is very difficult to associate ear image with a given person (in fact, most of users are not able to recognize their own ear image). Therefore, ear image databases do not have to be as much secured as face databases, since the risk of attacks is much lower.

On the other hand, ear biometrics is not a natural way of identifying humans. In real life we do not look at people ears to recognize them. Our identification decision is rather based on faces, voice or gait. The reason is that people lack in vocabulary to describe ears. The main task of ear biometrics is to define such vocabulary – in context of the computer vision

© 2009 Springer Science+Business Media, LLC

systems, such vocabulary is called "features." In ear biometrics computer vision systems, the main task is to extract such features, that will describe human ears in a distinctive way.

Even though ear biometrics have not been implemented commercially so far, there are already many methods of feature extraction from ear images developed. In this paper our goal is to overview these approaches and methods. The summary of the research groups with the proposed approaches and methods is given in the Table 1.

2D Ear Biometrics

In this section various approaches to 2D ear biometrics are presented. Firstly, methods based on geometrical parameters are overviewed. Then the global approach to feature extraction from 2D ear images is surveyed.

Geometrical Approach to Feature Extraction

The first to explore the possibility of using ear as a biometric in a computer vision system were Burge and Burger [4, 5]. They presented the geometrical method based on building neighborhood graphs and Voronoi diagrams of the detected edges.

Additionally, Burge and Burger pointed out that ▶ thermal imaging may solve the problem of ear occlusion (mainly by hair). They proposed to use segmentation algorithm based on color and texture of ear thermogram (ear thermal image).

Choraś developed several methods of geometrical feature extraction from ear images [6–8]. The proposed "Geometrical Parameters Methods" had been motivated by actual procedures used in the police and forensic evidence search applications. In reality, procedures of handling ear evidence (earprints and/or ear photographs) are based on geometrical features such as size, width, height and earlobe topology [1].

Choraś developed and tested the following methods in order to extract distinctive geometrical features from human ear 2D images:

- Concentric circles based method – *CCM*
- Contour tracing method – *CTM*
- Angle-based contour representation method – *ABM*
- Triangle ratio method – *GPM*– *TRM*
- Shape ratio method – *GPM*– *SRM*

Moreover, in Choraś work the contour detection algorithm and the method of ear contour image processing in order to select the most meaningful contours have been developed. Ear pre-classification method based on the longest contour orientation have also been proposed.

Choraś' methods were tested in laboratory-conditions. His ear image database was created in the controlled environment. The most effective methods were *GPM* and *CTM*. The Receiver Operating Characteristic curves for each of the geometrical method are presented in Fig. 1.

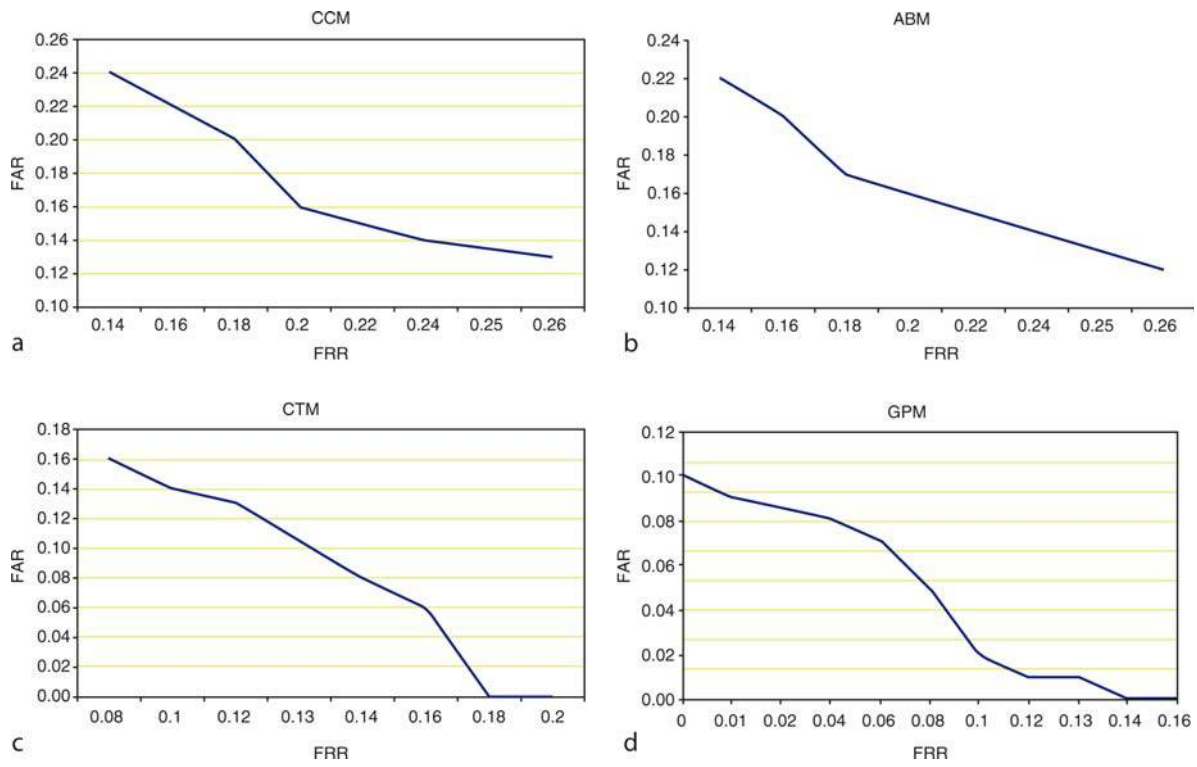
Yuan and Tian presented ear contour detection algorithm based on local approach [9]. Edge Tracking is applied to three regions in which contours were extracted in order to obtain clear, connected and non-disturbed contour, which may be further used in the recognition step.

Ear Biometrics. Table 1 Feature extraction approaches for ear biometrics

Research group	Proposed methodology
Burge and Burger	2D – Voronoi Diagrams
Choraś	2D – Geometrical Methods
Hurley et al.	2D – Force Field Transformation
Victor et al.	2D – PCA
Lu et al.	2D – ASM
Moreno et al.	2D – Compression Networks
Sana et al.	2D – Haar Wavelets
Yuan and Mu	2D – ASM
Arbab-Zavar	2D – SIFT, model
Chen and Bhanu	3D – ICP and shape descriptors
Yan and Bowyer	3D – ICP, edge-based and PCA
Cadavid and Abdel-Mottaleb	3D – SFM, SFS

Global Approach to Feature Extraction

Hereby, the global approaches to feature extraction from 2D ear images are presented. Principal Component Analysis, Force Field Transformations and



Ear Biometrics. **Figure 1** The Receiver Operating Characteristic (ROC) curve describing CCM, ABM, CTM and GPM methods (from left to right) [8].

Wavelets have been applied to ear biometrics human identification. Recently, the idea of recognition based on ear models gained some popularity and attention.

Victor et al. used Principal Component Analysis (*PCA*) in the experiment comparing ear and face properties in order to successfully identify humans in various conditions [10, 11].

In case of faces, the authors perform recognition on the basis of *eigen faces*. In case of ear biometrics, the authors used a set of *eigenears*. Their work proved that ear images are a very suitable source of data for identification and their results for ear images were not significantly different from those achieved for face images.

The proposed methodology, however, was not fully automated, since the reference (so called *landmarkpoints*) had to be manually inserted into images. In case of ear images these landmark points are manually marked in the Triangular Fossa and in the point known as Antitragus.

The sample ear image with the marked landmark points and the corresponding eigenear vector are shown in Fig. 2.

Hurley et al. introduced a method based on energy features of the 2D image [12, 13]. They proposed to perform force field transformation (step 1) in order to find energy lines, channels and wells (step 2).

Moreno et al. presented another approach to ear image feature extraction [14]. Their work was based on macrofeatures extracted by compression networks. Several neural networks methods and classifiers based on 2D intensity images were presented:

- Compression Networks,
- Borda Combination,
- Bayesian,
- Weighted Bayesian Combination.

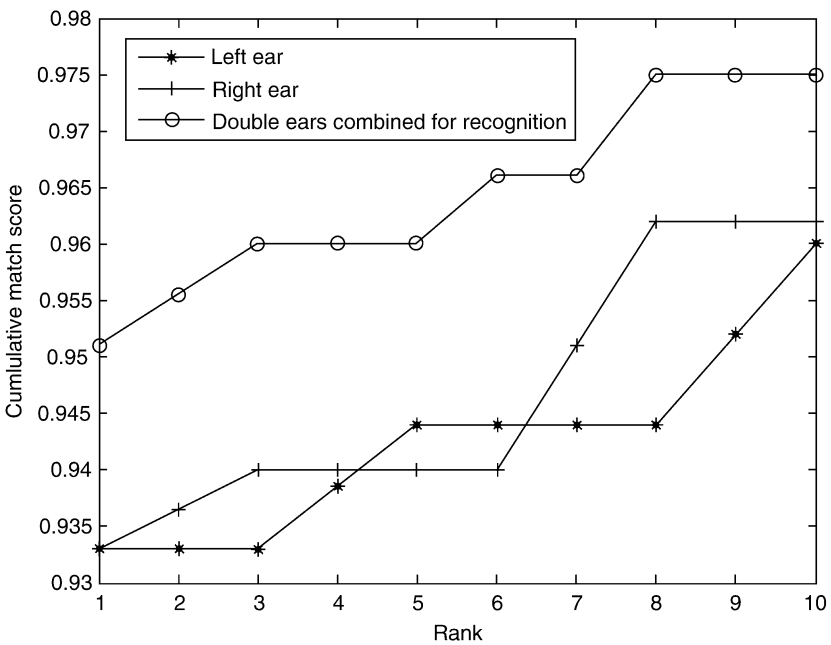
The best results of 93% were achieved by the Compression Network ear identification method.

Sana et al. developed a new approach to ear biometrics based on Haar wavelets [15]. After ear detection step, Haar Wavelet Transformation is applied and wavelet coefficients are computed. They performed their experiments on two ear datasets (from Indian Institute of Technology Kanpur and from Saugor University) and report accuracy of about 96% on both databases.

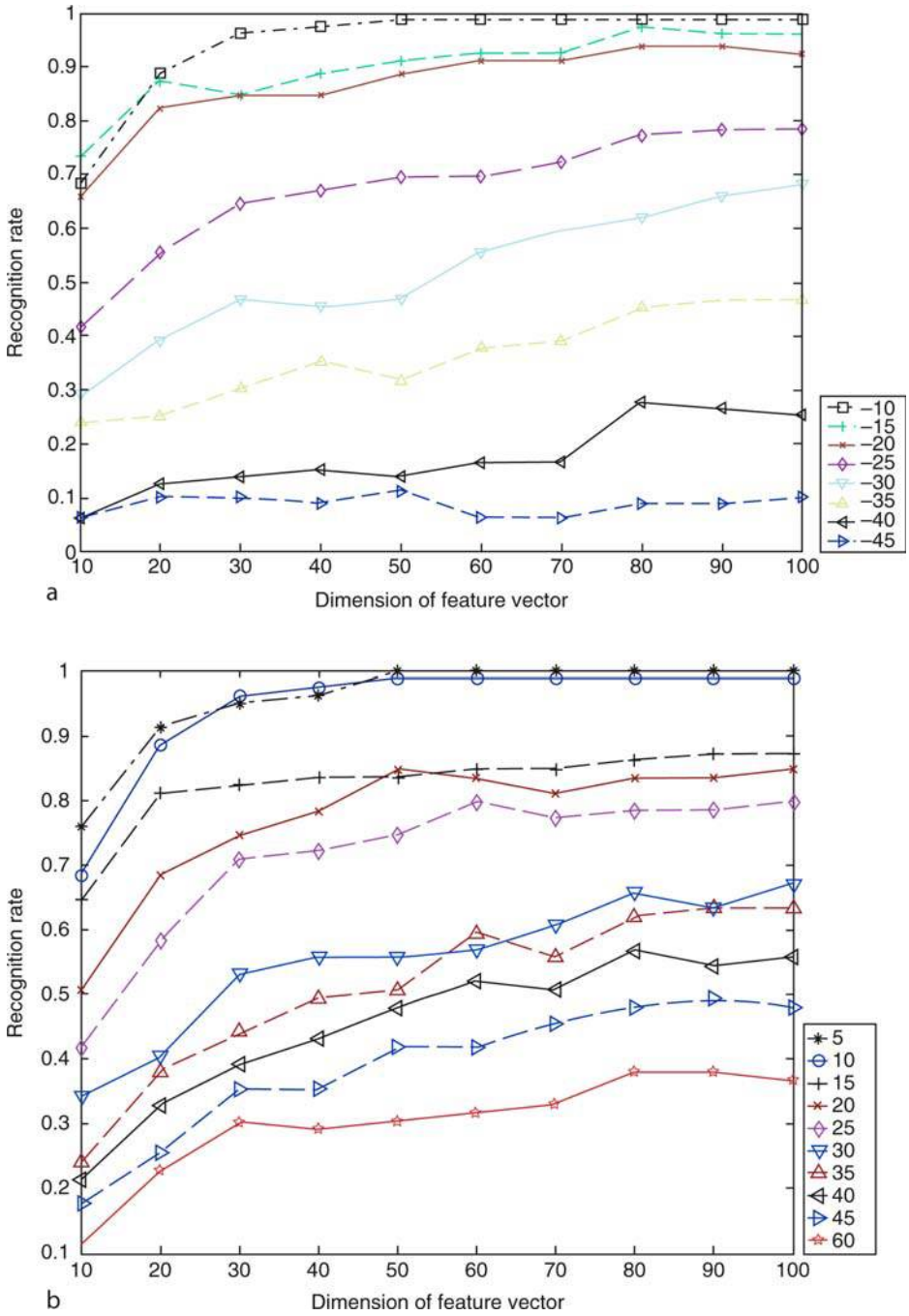
Lu et al. used Active Shape Models (*ASM*) to model the shape and local appearances of the ear in a statistical form [16]. Then *Eigenears* have been also used in a final classification step. They used both left and right ear images and showed that their fusion outperforms results achieved for single ears separately. They achieved 95.1% recognition rate for double ears. Their results for left, right and combined ears are presented in Fig. 3.

Yuan and Mu also explored the advantages of improved Active Shape Models (*ASM*) to the task of ear recognition [17]. They applied their algorithm to the rotation-invariance experiment. The interesting contribution of their work is the comparison of right and left rotation of the same ears. They found out that right head rotation of 20 degree is acceptable for recognition. For left head rotation, the acceptable angle is

Ear Biometrics. Figure 2 Ear image with the manually marked landmark points (Triangular Fossa and Antitragus) used in the PCA-based ear recognition [10].



Ear Biometrics. Figure 3 Cumulative Matching Score Curve for left, right and combined ears achieved by Lu et al. [16].

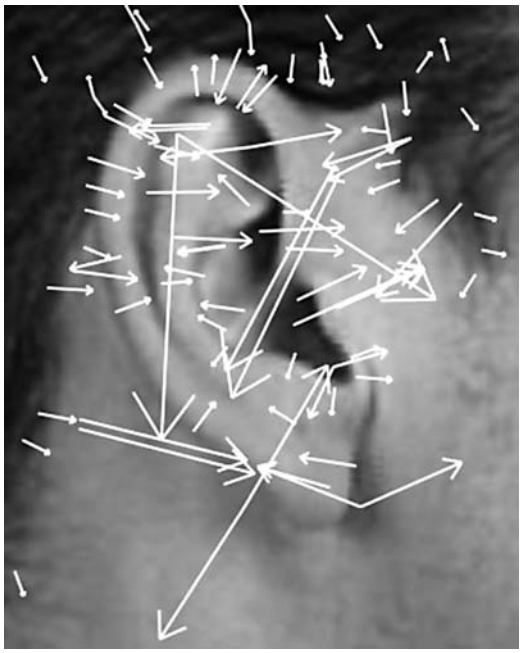


Ear Biometrics. Figure 4 Ear recognition for different degrees of left and right head rotations [17].

10 degree. Ear recognition results for different degrees of head rotations are presented in Fig. 4.

Arbab-Zavar et al. proposed to use Scale Invariant Feature Transform (*SIFT*) to extract the ear salient points and to create human ear model later used in recognition (Fig. 5) [18].

Their ear model is constructed using a stochastic method. In their experiments they proved that using ear model outperforms *PCA* method in case of occluded ears. The results of ear recognition for occluded ears (40% from top, and 40% from left) in comparison to *PCA* are given in Fig. 6.



Ear Biometrics. **Figure 5** Detected ear SIFT keypoints [18].

3D Ear Recognition

Recently, the possibility of human identification on the basis of 3D images have been extensively researched. Various approaches towards multimodal 2D+3D ear biometrics as well as 3D ear biometrics, mainly based on *ICP* (Iterative Closest Point), have been recently developed and published [19–21, 23].

Chen and Bhanu proposed 3D ear recognition based on local shape descriptor as well as two-step *ICP* algorithm [19]. Additionally, they developed the algorithm to detect ear regions from 3D range images. They collected their own ear image database (UCR database) consisting of 902 images from 302 subjects. Their results of ear detection, matching and identification are close to 100% recognition rate [20].

Yan and Bowyer developed three approaches to 3D ear recognition problem: edge-based, *ICP* and $3D-PCA$. Moreover they tested various approaches (for example 2D+3D) in multimodal biometric scenario [22].

They designed fully automated ear recognition system and achieved satisfactory results of 97.6% Rank-1 recognition [23]. The ear recognition performance results drawn as Receiver Operating Characteristic Curve and Cumulative Matching Score Curve are presented in Fig. 7 [23].

In their research they did not exclude partially occluded ears or ears with earrings. They performed experiments on the largest ear database collected so far. UND ear database is now becoming a standard ear database for ear recognition experiments. It is available for free at: www.nd.edu%/7Ecvrl/UNDBiometrics-Database.html.

Cadavid and Abdel-Mottaleb built 3D ear models from captured video frames. Then they used "structure from motion" (*SFM*) and "shape from shading" (*SFS*) techniques to extract 3D ear characteristics [24]. They were first to explore the 3D ear biometrics based on video sequences, not on images acquired by 3D range scanners.

Conclusion

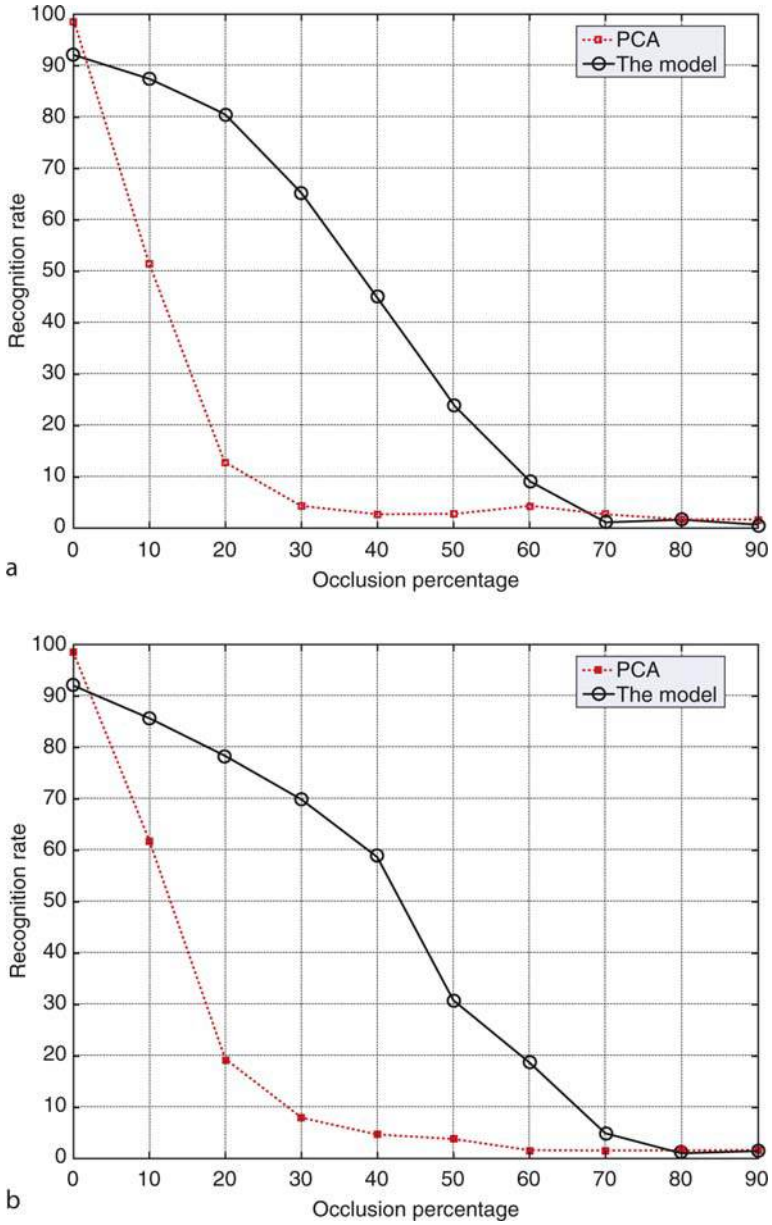
Human ear is a perfect source of data for passive person identification in many applications. In a growing need for security in various public places, ear biometrics seems to be a good solution, since ears are visible and their images can be easily taken, even without the examined person's knowledge.

The article presented an overview of various approaches and solutions to a problem of feature extraction from ear images. The summary of the research groups with the proposed approaches and methods is given in the Table 1.

It is noticeable that even though all of the proposed techniques are developed to solve the same image processing task, many totally different methodologies and algorithms have been developed.

Such situation proves that ear biometrics has lately gained much interest and popularity in computer science community. It also may be the indication that ear biometrics will become one of a standard means of human identification in unimodal or hybrid biometrics systems.

Ear biometrics can also be used to enhance effectiveness of other well-known biometrics, by its implementation in multimodal systems. Since most of the methods have some drawbacks, the idea of building multimodal (hybrid) biometrics systems is gaining lot of attention [25]. Due to its advantages, ear biometrics seems to be a good choice to support well known methods like voice, hand, palm or face identification.



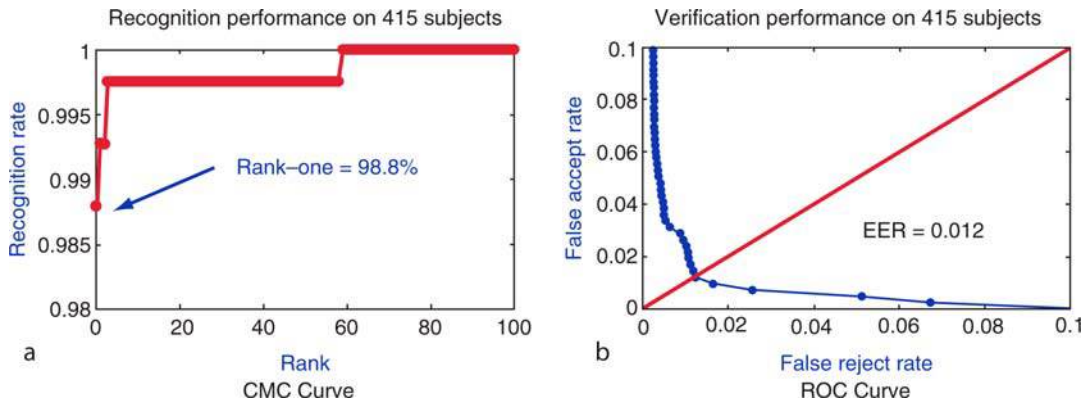
Ear Biometrics. **Figure 6** Ear recognition rates for occluded ear images, 40 % from top and 40% from left, respectively [18].

Summary

In this paper the holistic overview of ear recognition methods for biometrics applications is presented. 2D and 3D image processing algorithms applied to ear feature extraction are surveyed. Even though ear biometrics have not been implemented commercially so far, as pointed out by Hurley et al. ear biometrics is no longer in its infancy and has shown encouraging progress [26]. In this work strong motivation for using the ear as a biometrics is given, and afterwards, the geometrical approach to 2D ear biometrics, the global approach to 2D ear biometrics and 3D ear biometrics methods are presented, respectively.

Related Entries

► Ear Biometrics, 3D
► Physical Analogies for Ear Recognition



Ear Biometrics. Figure 7 Yan and Bowyer's 3D ear recognition results: Receiver Operating Characteristic Curve and Cumulative Matching Score Curve, respectively [23].

References

1. Kasprzak, J.: Forensic Otoscopy (in Polish), University of Warmia and Mazury Press (2003)
2. Iannarelli, A.: Ear Identification, Forensic Identification Series, Paramont Publishing Company (1989)
3. Choraś, M.: Ear biometrics in passive human identification systems. In: Proceedings of Pattern Recognition in Information Society. pp. 169–175, INSTICC, Paphos (2006)
4. Burge, M., Burger, W.: Ear biometrics. In: Jain, A.K., Bolle, R., Pankanti, S. (eds.) Biometrics: Personal Identification in Networked Society. pp. 273–286 (1998)
5. Burge, M., Burger, W.: Ear biometrics for machine vision. In: 15th International Conference of Pattern Recognition, pp. 826–830 IEEE (2000)
6. Choraś, M.: Ear biometrics based on geometrical feature extraction. J. ELCVIA (Comput. Vis. Image Analy.) **5**(3), 84–95 (2005)
7. Choraś, M.: Further developments in geometrical algorithms for ear biometrics. In: Proceedings of Articulated Motion and Deformable Objects – AMDO 2006, pp. 58–67. LNCS 4069, Springer Berlin (2006)
8. Choraś, M.: Perspective methods human identification: ear biometrics. Opto-Electr. Rev. **16**(1), 49–60 (2008)
9. Yuan, W. Tian, Y.: Ear Contour Detection based on Edge Tracking. In: Proceedings of Intelligent Control and Automation. pp. 10450–10453 IEEE, Dalian, China (2006)
10. Victor, B., Bowyer, K.W., Sarkar, S.: An evaluation of face and ear biometrics. In: Proceedings of International Conference on Pattern Recognition, pp. 429–432 (2002)
11. Chang, K., Victor, B., Bowyer, K.W., Sarkar, S.: Comparison and combination of ear and face images in appearance-based biometrics. IEEE Trans. Patt. Analy. Mach. Intell. **25**(8), 1160–1165 (2003)
12. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field energy functionals for image feature extraction. Image Vis. Comput. J. **20**(5–6), 311–318 (2002)
13. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field energy functionals for ear biometrics. Comput. Vis. Image Understand **98**(3), 491–512 (2005)
14. Moreno, B., Sanchez, A., Velez, J.F.: On the use of outer ear images for personal identification in security applications. In: Proceedings of IEEE Conference on Security Technology, pp. 469–476 (1999)
15. Sana, A., Gupta, P., Purkai, R.: Ear biometrics: a new approach. In: Pal, P. (ed.) Advances in Pattern Recognition, pp. 46–50, World Scientific Publishing (2007)
16. Lu, L., Zhang, X., Zhao, Y., Jia, Y.: Ear recognition based on statistical shape model. In: Proceedings of International Conference on Innovative Computing, Information and Control, vol. 3, pp. 353–356. IEEE (2006)
17. Yuan, L., Mu, Z.: Ear recognition based on 2D images. In: Proceedings of IEEE Conference on Biometrics: Theory, Applications and Systems – BTAS'07. Washington (2007)
18. Arab-Zavar, B., Nixon, M.S., Hurley, D.J.: On model-based analysis of ear biometrics. In: Proceedings of IEEE Conference on Biometrics: Theory, Applications and Systems – BTAS'07. Washington (2007)
19. Chen, H., Bhanu, B.: Contour matching for 3D ear recognition. In: Proceedings of Workshop on Applications of Computer Vision (WACV), pp. 123–128 (2005)
20. Chen, H., Bhanu, B.: Human ear recognition in 3D. IEEE Trans. Pattern Analy. Mach. Intell. **29**(4), 718–737 (2007)
21. Yan, P., Bowyer, K.W.: ICP-based approaches for 3D ear recognition. In: Proceedings of SPIE Biometric Technology for Human Identification, pp. 282–291 (2005)
22. Yan, P., Bowyer, K.W.: Multi-biometrics 2D and 3D ear recognition. In: Proceedings of Audio- and Video-based Biometric Person Authentication, pp. 503–512 (2005)
23. Yan, P.: Ear Biometrics in Human Identification. Dissertation, University of Notre Dame (2006)
24. Cadavid, S., Abdel-Mottaleb, M.: Human identification based on 3D ear models. In: Proceedings of IEEE Conference on Biometrics: Theory, Applications and Systems – BTAS'07, Washington (2007)
25. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. International Series on Biometrics, Springer, Berlin (2006)
26. Hurley, D.J., Arab-Zavar, B., Nixon, M.S.: The ear as a biometric. In: Proceedings of Eusipco'07, pp. 25–29 Poznan (2007)

Ear Biometrics, 3D

Bir Bhanu, Hui Chen
Center for Research in Intelligent Systems, University of California, Riverside, CA, USA

Synonym

Ear Recognition, 3D

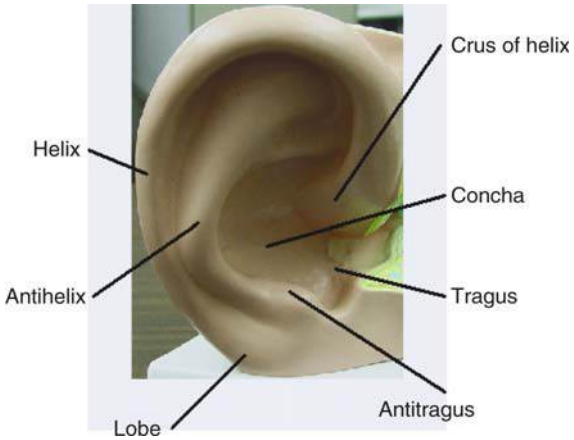
Definition

The human ear is a new class of relatively stable biometrics. After decades of research of anthropometric measurements of ear photographs of thousands of people, it has been found that no two ears are alike, even in the cases of identical and fraternal twins, triplets, and quadruplets [1]. It is also found that the structure of the ear does not change radically over time. Ear biometric has played a significant role in forensic science and its use by law enforcement agencies for many years [1] but most of this work has been on analyzing the ▶ earprints manually. Recent work on ear biometrics focuses on developing automated techniques for ear recognition [2]. Ear biometrics can be based on a 2D gray scale or color image, 3D range image, or a combination of 2D and 3D images. Typically, an ear biometric system consists of ear detection and ear recognition modules.

Introduction

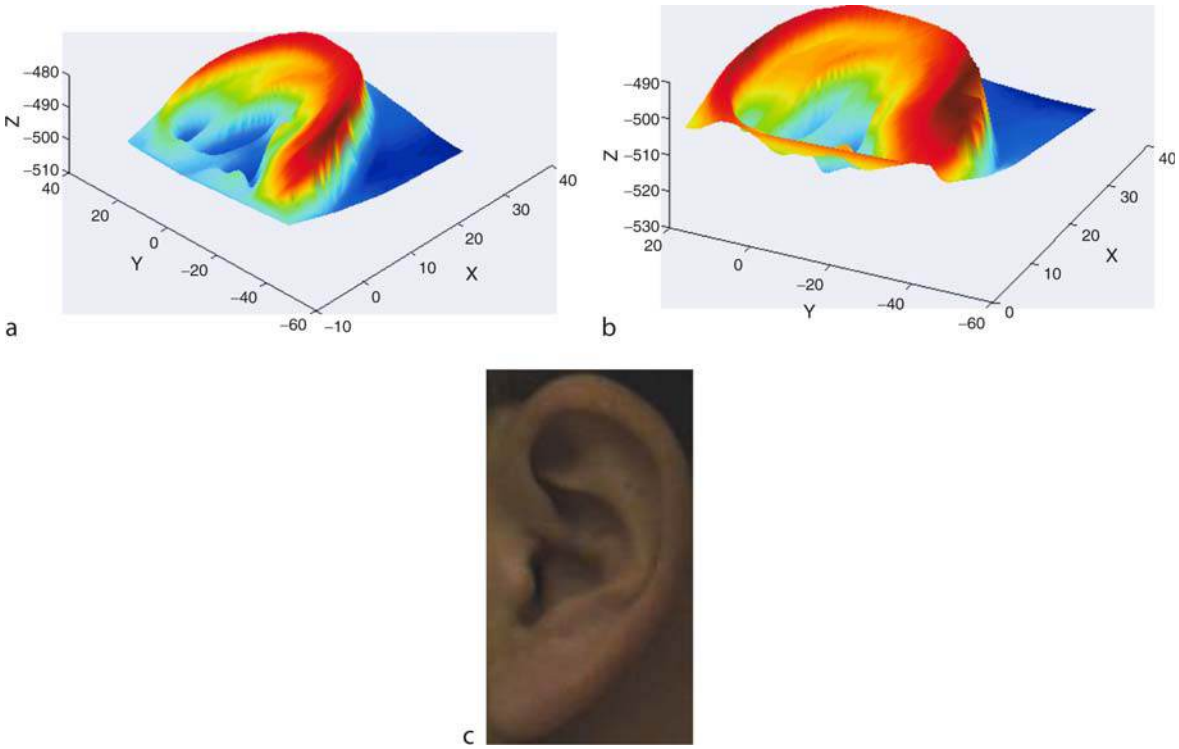
Rich in features, the human ear is a stable structure that does not change much in shape with the age and with facial expressions (see Fig. 1). Ear can be easily captured from a distance without a fully cooperative subject although it can sometimes be hidden by hair, muffler, scarf, and earrings. Researchers have developed several biometric techniques using the 2D intensity images of human ears [3–8].

Burge and Burger [3, 4] developed a computer vision system to recognize ears in the intensity images. Their algorithm consisted of four components: edge extraction, curve extraction, construction of a graph model from the Voronoi diagram of the edge segments, and graph matching. Hurley et al. [5] applied a force field transform to the entire ear image and extracted wells and channels. The wells and channels form the basis of an ear's signature. To evaluate differences among ears, they used a measure of the average normalized distance of the well positions, together with the accumulated direction to the position of each well point from a chosen reference point. Later, Hurley et al. [6] measured convergence to achieve greater potency in recognition. Chang et al. [8] used principal component analysis for ear and face images and performed experiments with face, ear, and face plus ear. Their results showed that multi-modal recognition using both face and ear achieved a much better performance than the individual biometrics.



Ear Biometrics, 3D. Figure 1 The external ear and its anatomical parts.

The performance of these 2D techniques is greatly affected by the pose variation and imaging conditions. However, ear can be imaged in 3D using a range sensor which provides a registered color and range image. Figure 2 shows an example of a range image and the registered color image acquired by a Minolta Vivid 300 camera. A range image is relatively insensitive to illuminations and contains surface shape information related to the anatomical structure, which makes it possible to develop a robust 3D ear biometrics. Examples of ear recognition using 3D data are [9–13]. The performance of 3D approaches for ear recognition is significantly higher than the 2D approaches. In the following, the chapter focuses on 3D approaches for ear detection and recognition.



Ear Biometrics, 3D. Figure 2 Range image and color image captured by a Minolta Vivid 300 camera. In images (**a**) and (**b**), the range image of one ear is displayed as the shaded mesh from two viewpoints (the units of x, y and z are in millimeters). Image (**c**) shows the color image of the ear.

Datasets

There are currently two datasets for 3D ear performance evaluation: The University of California at Riverside dataset (the UCR dataset) and the University of Notre Dame public dataset (the UND dataset). In the UCR dataset there is no time lapse between the gallery and probe for the same subject, while there is a time lapse of a few weeks (on the average) in the UND dataset.

UCR Dataset: The data [10] are captured by a Minolta Vivid 300 camera. This camera uses the light-stripe method to emit a horizontal stripe light to the object and the reflected light is then converted by triangulation into distance information. The camera outputs a range image and its registered color image in less than 1 s. The range image contains 200 × 200 grid points and each grid point has a 3D coordinate (x, y, z) and a set of color (r, g, b) values. During the acquisition, 155 subjects sit on a chair about 0.55–0.75 m from the camera in an indoor office environment. The first shot is taken when a subject's left-side face is approximately parallel to the image plane; two shots are taken when the subject is asked to rotate his or her head to the left and to the right side within $\pm 35^\circ$ with respect to his or her torso. During this process, there can be some face tilt as well, which is not measured. A total of six images per subject are recorded. A total of 902 shots are used for the experiments since some shots are not properly recorded. Every person has at least four shots. The average number of points on the side face scans is 23,205. There are there different poses in the collected data: frontal, left, and right. Among the total 155 subjects, there are 17 females. Among the 155 subjects, 6 subjects have earrings and 12 subjects have their ears partially occluded by hair (with less than 10% occlusion).

UND Dataset: The data [13] are acquired with a Minolta Vivid 910 camera. The camera outputs a 480 × 640 range image and its registered color image of the same size. During acquisition, the subject sits approximately 1.5 m away from the sensor with the left

side of the face toward the camera. In Collection F, there are 302 subjects with 302 time-lapse gallery-pro. Collection G contains 415 subjects of which 302 subjects are from Collection F. The most important part of Collection G is that it has 24 subjects with images taken at four different viewpoints.

Ear Detection

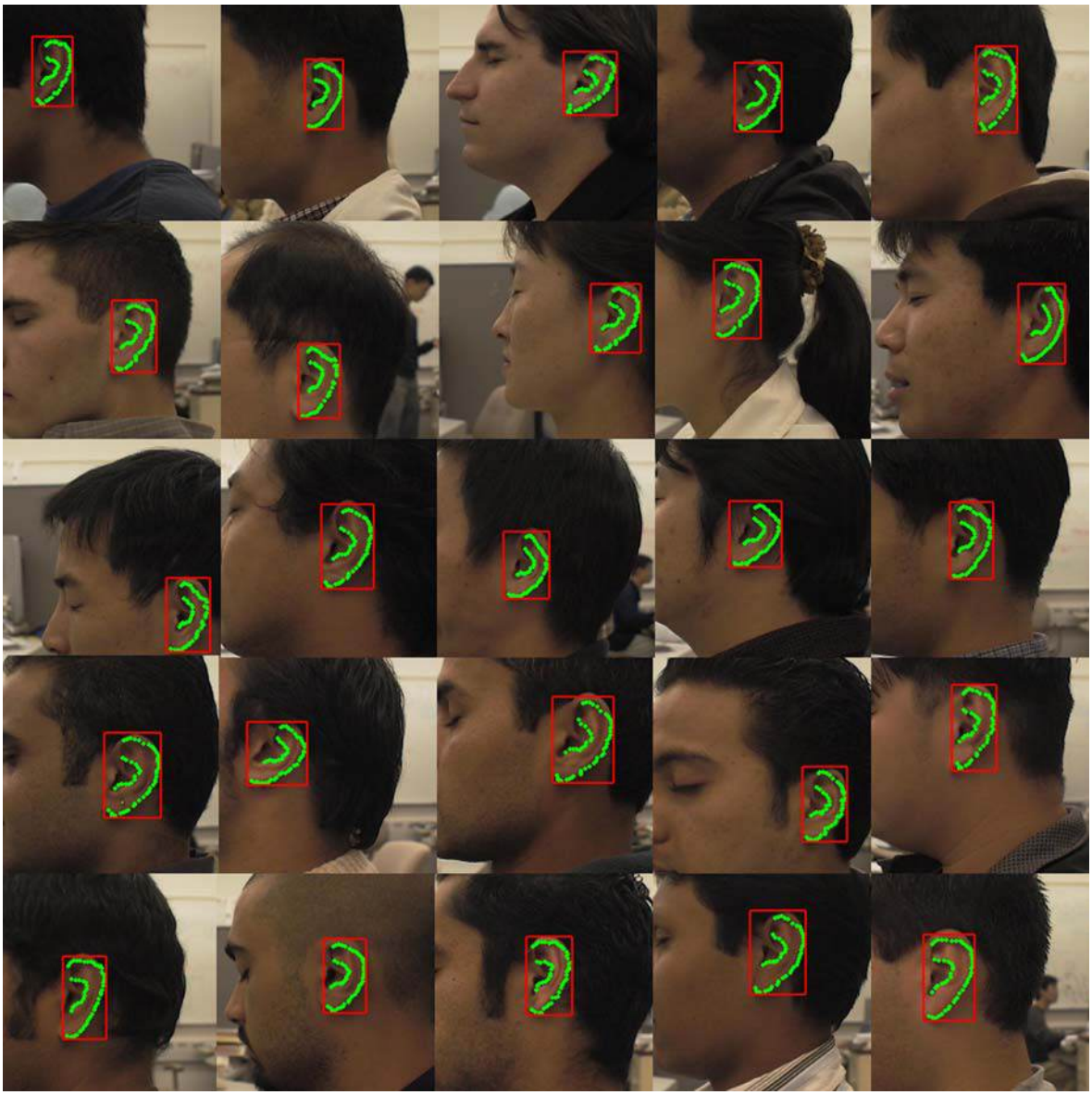
Human ear detection is the first task of a human ear recognition system and its performance significantly affects the overall quality of the system. Automated techniques for locating human ears in side face range images are: (i) template matching based detection, (ii) ear shape model based detection, and (iii) fusion of color and range images and global-to-local registration based detection. The first two approaches use range images only, and the third approach fuses the color and range images.

The template matching based approach has two stages: offline model template building and online ear detection. The ear can be thought of as a rigid object with much concave and convex areas. The averaged histogram of ▶ shape index (a quantitative measure of the shape of a surface) represents the ear model template. During the online detection, first the step edges are computated and thresholded since there is a sharp step edge around the ear boundary, and then image dilation and connected-component analysis is performed to find the potential regions containing an ear. Next, for every potential region, the regions are grown and the dissimilarity between each region's histogram of shape indexes and the model template is computed. Finally, among all of the regions, we choose the one with the minimum dissimilarity as the detected region that contains ear.

For the ear shape model based approach, the ear shape model is represented by a set of discrete 3D vertices corresponding to ear helix and anti-helix parts. Since the two curves formed by the ear helix and anti-helix parts are similar for different people, we do not take into account the small deformation of two curves between different persons, which greatly simplifies the ear shape model. Given side face range images, first the step edges are extracted; then the edge segments are dilated, thinned, and grouped into different clusters which are the potential regions containing an ear. For each cluster, the ear shape model is registered with the edges. The region with the minimum mean registration error is declared as the detected ear region; the ear helix and anti-helix parts are identified in this process.

In the above two approaches, there are some edge segments caused by non-skin pixels, which result in the false detection. Since a range sensor provides a registered 3D range image and a 2D color image (see Fig. 2), it is possible to achieve a better detection performance by fusion of the color and range images. This approach consists of two-steps for locating the ear helix and the anti-helix parts.

In the first step a skin color classifier is used to isolate the side face in an image by modeling the skin color and non-skin color distributions as a mixture of Gaussians. The edges from the 2D color image are combined with the step edges from the range image to locate regions-of-interest (ROIs) that may contain an ear. In the second step, to locate an ear accurately, the reference 3D ear shape model, which is represented by a set of discrete 3D vertices on the ear helix and the anti-helix parts, is adapted to individual ear images by following a global-to-local registration procedure instead of training an active shape model built from a large set of ears to learn the shape variation. In this procedure after the initial global registration local deformation process is carried out where it is necessary to preserve the structure of the reference ear shape model since neighboring points cannot move independently under the deformation due to physical constraints. The bending energy of thin plate spline, a quantitative measure for non-rigid deformations, is incorporated into the optimization formulation as a regularization term to preserve the topology of the ear shape model under the shape deformation. The optimization procedure drives the initial global registration toward the ear helix and the anti-helix parts, which results in the one-to-one correspondence of the ear helix and the anti-helix between the reference ear shape model and the input image. Figure 3 shows various examples in which the detected ear helix and the anti-helix parts are shown by the dots superimposed on the 2D color images and the detected ear is bounded by the rectangular box. We observe that the ears and their helix and anti-helix parts are correctly detected. This approach provides very high detection accuracy. A comparison of the three approaches shows that the first approach runs the fastest and it is simple, effective, and easy to implement. The second approach locates an ear more

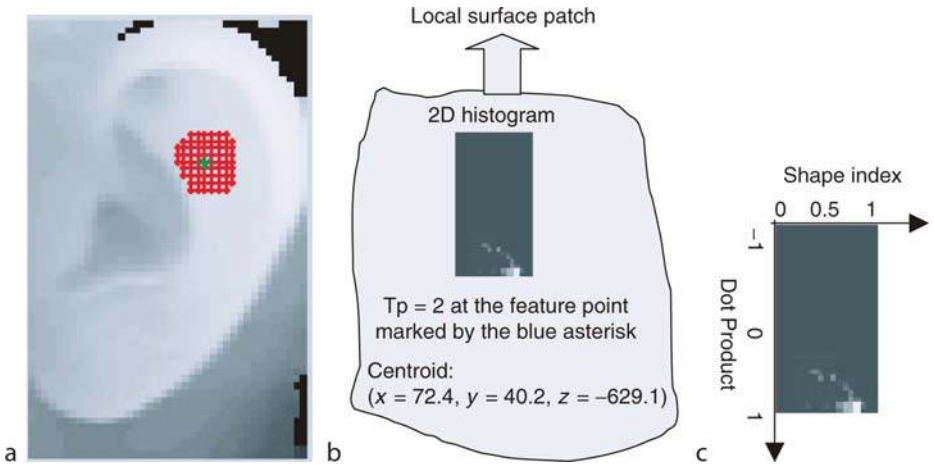


Ear Biometrics, 3D. **Figure 3** Results of ear localization on the UCR dataset. The helix and the anti-helix parts are marked by the bright dots and the detected ear is bounded by a rectangular box.

accurately than the first approach since the shape model is used. The third approach performs the best on both the UCR and the UND datasets and it runs slightly slower than the other approaches.

Ear Recognition

The approach for ear detection is followed to build a database of ears that belong to different people. For ear recognition, two representations are used: the ear helix/ antihelix representation obtained from the detection algorithm and a new ▶ local surface patch representation computed at feature points to estimate the initial rigid transformation between a gallery-probe pair. For the ear helix/antihelix representation, the correspondence of ear helix and antihelix parts (available from the ear detection algorithm) between a gallery-probe ear pair is established and it is used to compute the initial rigid transformation. For the local surface patch (LSP) representation, a local surface descriptor (see Fig. 4) is characterized by a centroid, a



Ear Biometrics, 3D. **Figure 4** Illustration of a local surface patch (LSP). (**a**) Feature point *P* is marked by the asterisk and its neighbors *N* are marked by the interconnected dots. (**b**) LSP representation includes a 2D histogram, a surface type and centroid coordinates. (**c**) The 2D histogram is shown as a gray image in which the brighter areas correspond to bins with the high frequency of occurrence.

local surface type, and a 2D histogram. The 2D histogram and surface type are used for comparison of LSPs and the centroid is used for computing the rigid transformation. The patch encodes the geometric information of a local surface. The local surface descriptors are computed for the feature points, which are defined as either the local minimum or the local maximum of shape indexes. By comparing the local surface patches for a gallery and a probe image, the potential corresponding local surface patches are established and then filtered by geometric constraints. Based on the filtered correspondences, the initial rigid transformation is estimated. Once this transformation is obtained using either of the two representations, it is then applied to randomly selected control points of the hypothesized gallery ear in the database. A modified iterative closest point (ICP) (▶ ICP algorithm) algorithm is run to improve the transformation, which brings a gallery ear and a probe ear into the best alignment, for every gallery probe pair. The root mean square (RMS) registration error is used as the matching error criterion. The subject in the gallery with the minimum RMS error is declared as the recognized person in the probe.

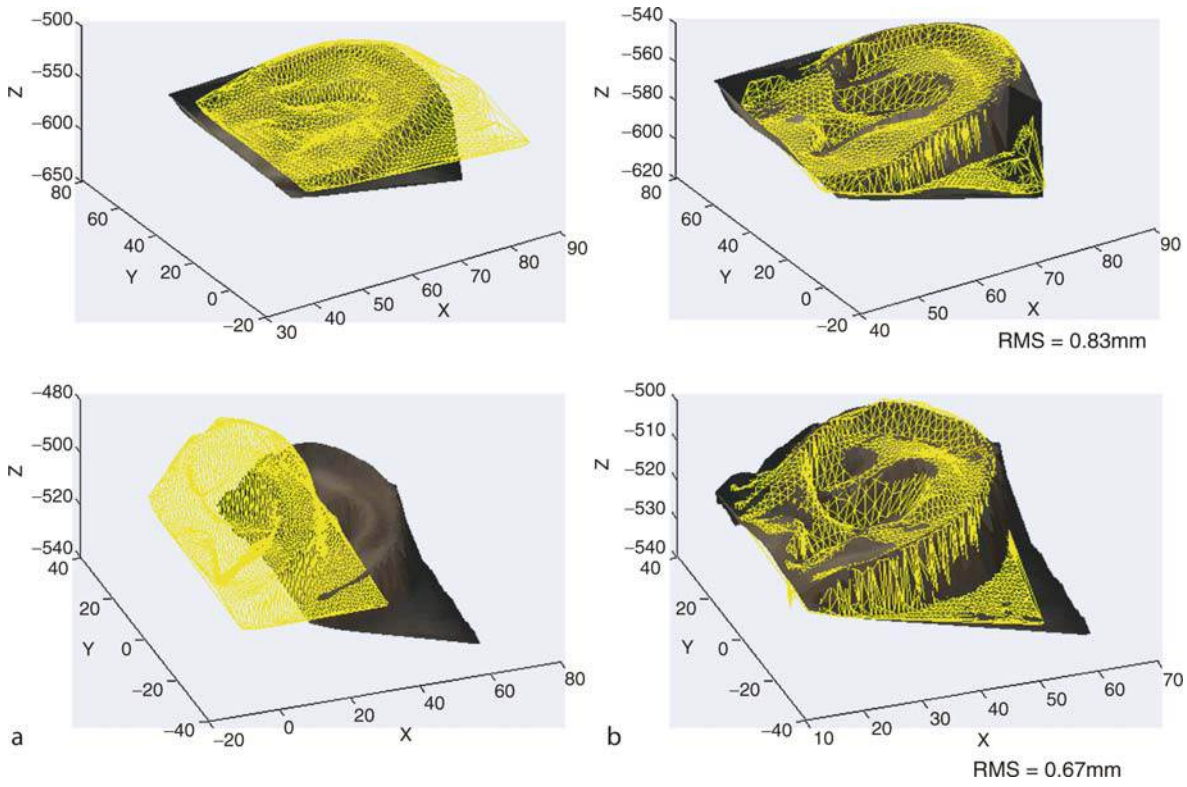
The experiments are performed on the the UCR data set and the UND data.

Examples of correctly recognized gallery-probe ear pairs using the helix/anti-helix representation is shown in Fig. 5. Similarly, examples of correctly recognized gallery-probe ear pairs using local surface patch representation are shown in Fig. 6. From Figs. 5 and 6, we observe that each gallery ear is well aligned with the corresponding probe ear.

The recognition results are shown in Table 1. In order to evaluate the proposed surface matching schemes, we perform experiments under two scenarios: (1) One frontal ear of a subject is in the gallery set and another frontal ear of the same subject is in the probe set and (2) Two frontal ears of a subject are in the gallery set and the rest of the ear images of the same subject are in the probe set. These two scenarios are denoted as ES1 and ES2, respectively. ES1 is used for testing the performance of the system to recognize ears with the same pose; ES2 is used for testing the performance of the system to recognize ears with pose variations.

A comparison of the LSP representation with the spin image representation for identification and verification is given in [10]. This comparison showed that the LSP representation achieved a slightly better performance than the spin image representation.

For the identification, usually a biometrics system conducts a one-to-many comparison to establish an individual's identity. This process is computationally expensive, especially for a large database. There is a need to develop a general framework for rapid recognition of 3D ears. An approach that combines the feature embedding and support vector machine (SVM) rank learning techniques is described in [2]. It provides a sublinear



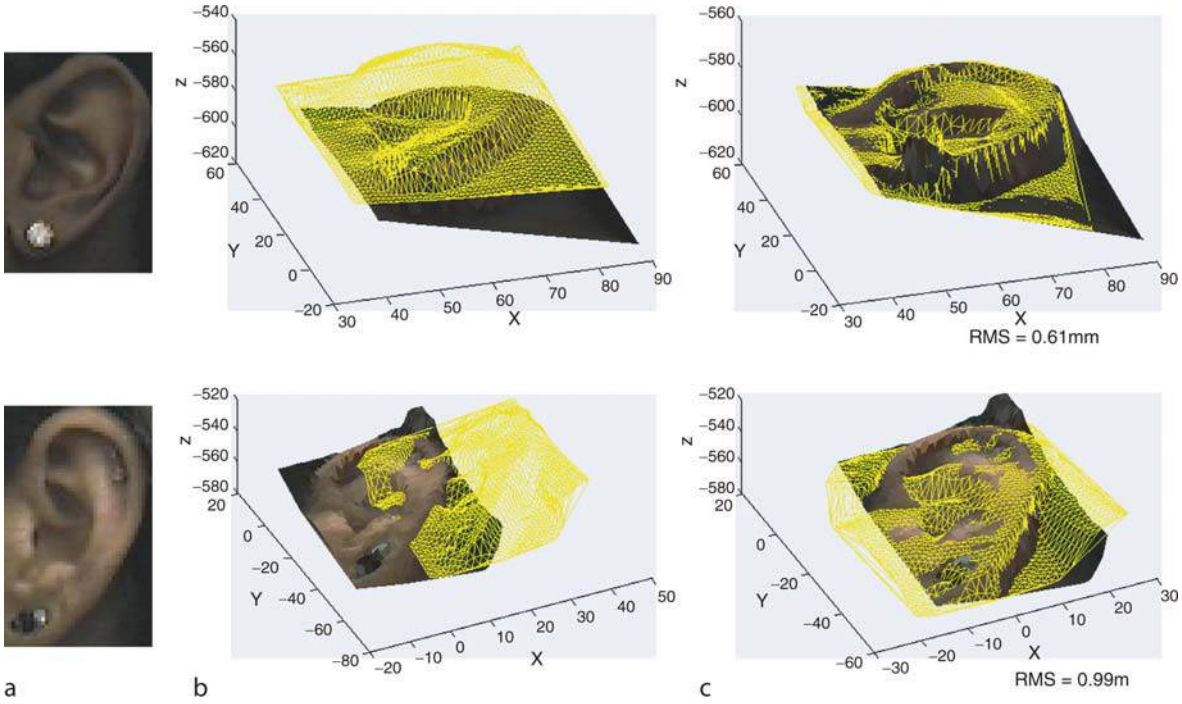
Ear Biometrics, 3D. **Figure 5** Two examples of correctly recognized gallery-probe pairs using the ear helix/anti-helix representation. (**a**) Examples of probe ears with the corresponding gallery ears before alignment. (**b**) Examples of probe ears with the correctly recognized gallery ears after alignment. The gallery ear represented by the mesh is overlaid on the textured 3D probe ear. The units of x, y and z are millimeters (mm).

time complexity on the number of models without making any assumptions about the feature distributions. The experimental results on the UCR dataset (155 subjects with 902 ear images) and the UND dataset (302 subjects with 604 ear images) containing 3D ear objects demonstrated the performance and effectiveness of the approach. The average processing time per query are 72 and 192 s, respectively, on the two datasets with the reduction by a factor of 6 compared with the sequential matching without feature embedding. With this speed-up, the recognition performances on the two datasets degraded 5.8% and 2.4%, respectively. The performance of this algorithm is scalable with the database size without sacrificing much accuracy.

The prediction of the performance of a biometric system is also an important consideration in the real world applications. Match and non-match distances obtained from matching 3D ears are used to estimate their distributions. By modeling cumulative match characteristic (CMC) curve as a binomial distribution, the ear recognition performance can be predicted on a larger gallery [2]. The performance prediction model in [2] showed the scalability of the proposed ear biometrics system with increased database size.

Summary

Ear recognition, especially in 3D, is a relatively new area in biometrics research.The experimental results on the two large datasets show that ear biometrics has the potential to be used in the real-world applications to identify/authenticate humans by their ears. Ear biometrics can be used in both the low and high security applications and in combination with other biometrics such as face. With the decreasing



Ear Biometrics, 3D. **Figure 6** Two examples of the correctly recognized gallery-probe pairs using the LSP representation. The ears have earrings. Images in column (**a**) show color images of ears. Images in column (**b**) and (**c**) show the probe ear with the corresponding gallery ear before the alignment and after the alignment, respectively. The gallery ears represented by the mesh are overlaid on the textured 3D probe ears. The units of x, y and z are in millimeters (mm).

Ear Biometrics, 3D. **Table 1** Recognition results on UCR and UND datasets using helix/anti-helix and LSP representation

Dataset	Helix/anti-helix representation					LSP representation				
	Rank-1	Rank-2	Rank-3	Rank-4	Rank-5	Rank-1	Rank-2	Rank-3	Rank-4	Rank-5
UCR ES_1(155,155)	96.77%	98.06%	98.71%	98.71%	98.71%	94.84%	96.77%	96.77%	96.77%	96.77%
UCR ES_2(310,592)	94.43%	96.96%	97.80%	98.31%	98.31%	94.43%	96.96%	97.30%	97.64%	97.80%
UND(302,302)	96.03%	96.69%	97.35%	97.68%	98.01%	96.36%	98.01%	98.34%	98.34%	98.34%

cost and size of a 3D scanner and the increased performance, we believe that 3D ear biometrics will be highly useful in many real-world applications in the future. It is possible to use the infrared images of ears to overcome the problem of occlusion of the ear by hair. Recent work in acoustics allows one to (a) determine the impulse response of an ear [14] and (b) make use of otoacoustic emissions [15] as a biometric. Thus, it is possible to combine shape-based ear recognition with the acoustic recognition of ear to develop an extremely fool-proof system for recognizing a live individual.

Related Entries

- ▶ Face Recognition
- ▶ Face Recognition, Overview

References

1. Iannarelli, A.: Ear Identification. Forensic Identification Series. Paramont Publishing Company, (1989)
2. Bhanu, B., Chen, H.: Human Ear Recognition by Computer. Springer (2008)

3. Burge, M., Burger, W.: Ear biometrics. in A. Jain, R. Bolle, S. Pankanti, Biometrics - Personal Identification in Networked Society, Kluwer Academic Publishers (1999)
4. Burge, M., Burger, W.: Ear biometrics in computer vision. Proc. Int. Conf. on Pattern Recognition **2**, 822–826 (2000)
5. Hurley, D.J., Nixon, M., Carter, J.N.: Force field energy functionals for image feature extraction. Image and Vision Computing **20**(5–6), 311–317 (2002)
6. Hurley, D., Nixon, M., Carter, J.: Force field feature extraction for ear biometrics. Computer Vision and Image Understanding **98**(3), 491–512 (2005)
7. Hurley, D., Arbab-Zavar, B., Nixon, M.: The ear as a biometric. in A. Jain, P. Flynn, A. Ross, Handbook of Biometrics, Springer (2007)
8. Chang, K., Bowyer, K.W., Sarkar, S., Victor, B.: Comparison and combination of ear and face images in appearance-based biometrics. IEEE Trans. Pattern Analysis and Machine Intelligence **25**(9), 1160–1165 (2003)
9. Bhanu, B., Chen, H.: Human ear recognition in 3D. Proc. Workshop on Multimodal User Authentication pp. 91–98 (2003)
10. Chen, H., Bhanu, B.: Human ear recognition in 3D. IEEE Trans. Pattern Analysis and Machine Intelligence **29**(4), 718–737 (2007)
11. Chen, H., Bhanu, B.: 3D free-form object recognition in range images using local surface patches. Pattern Recognition Letters **28**(10), 1252–1262 (2007)
12. Yan, P., Bowyer, K.W.: Multi-biometrics 2D and 3D ear recognition. Proc. Audio and Video Based Biometric Person Authentication pp. 503–512 (2005)
13. Yan, P., Bowyer, K.W.: Biometric recognition using 3D ear shape. IEEE Trans. Pattern Analysis and Machine Intelligence **29**(8), 1297–1308 (2007)
14. Akkermans, A., Kevenaar, T., Schobben, D.: Automatic ear recognition for person identification. Proc. IEEE Workshop on Automatic Identiifcation Advanced Technologies pp. 219–223 (2005)
15. Swabey, M., Beeby, S.P., Brown, A.: Using otoacoustic emissions as a biometric. Proc. First International Conference on Biometric Authentication pp. 600–606 (2004)

Ear Recognition

► Ear Biometrics

Earmark(s)

Earmark(s) are the ear impression(s) recovered typically from the crime scene.

► Earprints, Forensic Evidence of

Earprints

Earprints are the control impressions taken from the ears of known potential donors.

► Earprints, Forensic Evidence of

Earprints, Forensic Evidence of

Christophe Champod
Institut de Police Scientifique, Ecole des Sciences Criminelles, Université de Lausanne, Switzerland

Synonyms

Earprints; Earmark(s); Identification; ACE-V

Definition

Forensic evidence of earprint is the field of forensic science devoted to the collection and comparison of ► earmark(s) (generally left in association to a crime scene) with earprints obtained from ears of individuals of interest under controlled condition. Anthropometric studies and empirical evidence have shown that the forms left by an ear are very discriminating and allow bringing evidence of reasonable strength regarding the identity of sources.

Current research aims at bringing structured data relevant to the forensic examination process and move from a field dominated by subjectively informed experience and anecdotal evidence to a field where transparent data allows an assessment of the case.

Introduction

The use of earmarks in forensic science is a consequence of the recovery of such marks during crime scene investigation. Earmarks are left on surfaces where one applies his or her ear to listen. The deposition mechanism is similar to the mechanism whereby

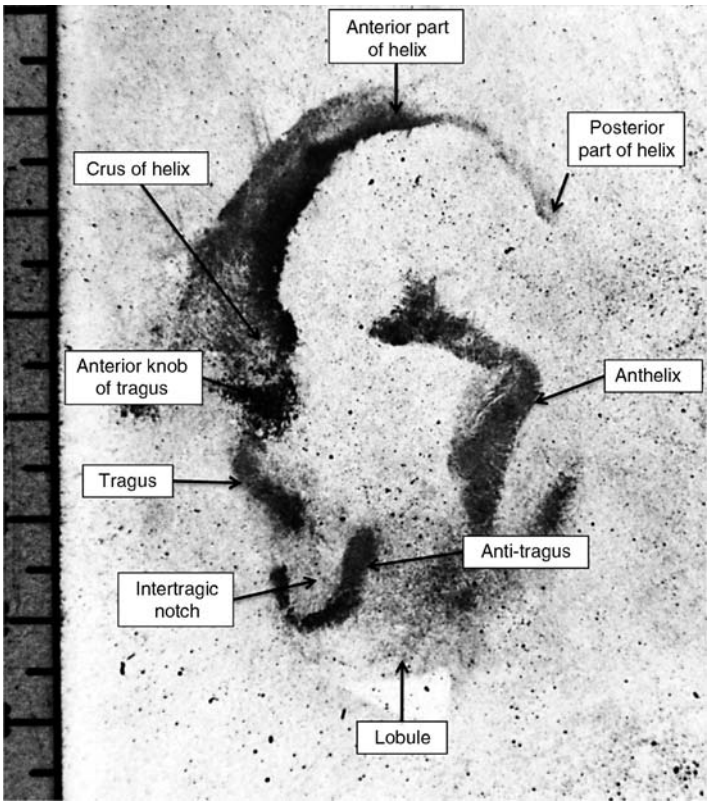
fingermarks are deposited on surfaces when touched with bare hands. Secretions originating mainly from sebaceous glands cover the ear. When the ear is put in contact with a surface it leaves a mark (often not readily visible), a form corresponding to the shape of the external organ applied. Such marks are often detected in conjunction with the search for latent fingermarks using the same detection techniques. The surface systematically searched for earmarks are the points of access (doors or windows) and their recovery translates a typical modus operandi for the perpetration of the offence. Marks are generally detected by applying a powdering technique on the surface. Then the mark is described, located and photographed, and preserved on an adhesive or gelatine lifter.

An example of a recovered mark is shown in Fig. 1 with indications of the typical nomenclature used to describe the features of the ear (the figure shows directly the mark, whereas these anatomical descriptions refer to the ear itself).

A useful model that helps scientists focus on their role is called the 'Investigator/Evaluator' dichotomy. In reality, scientists operate in both investigator and evaluator modes in many of the cases. Providing opinion in these two different modes requires different mindsets. An understanding of these differences is essential in the context of earprints analysis.

In *Investigator* mode, indeed it is the scientist's role to form a reasonable hypothesis from the observations. While attending a crime scene and recovering earmarks, the police may put forward the following investigative questions:

- How many people were involved?
- What potential set of actions may have given rise to this (these) mark(s)?
- What is the range of height of the person at the source of that (these) mark(s)?
- Using reference collections or databases, could you suggest a name to the investigation?



Earprints, Forensic Evidence of. **Figure 1** Earmark recovered from a windowpane. Anatomical features are designated with arrows.

The scientist will form and communicate what may explain the observations based on his knowledge, experience or through the use of databases. Generally, scientists operate in this mode before a suspect is arrested and charged with an offence. Opinions provide directions and options to the investigation and it is accepted that some directions offered may be misleading. The problem arises when this data is not further scrutinized and used as evaluative evidence in court. In *evaluator* mode, the role of the scientist is to form a view on the weight of evidence to be assigned to the scientific findings. This is the primary role of the scientist in what may be called *post-charge* cases; i.e., cases in which a suspect has been arrested and charged. In this role, the concept of weight of evidence associated with the findings should be approached more carefully.

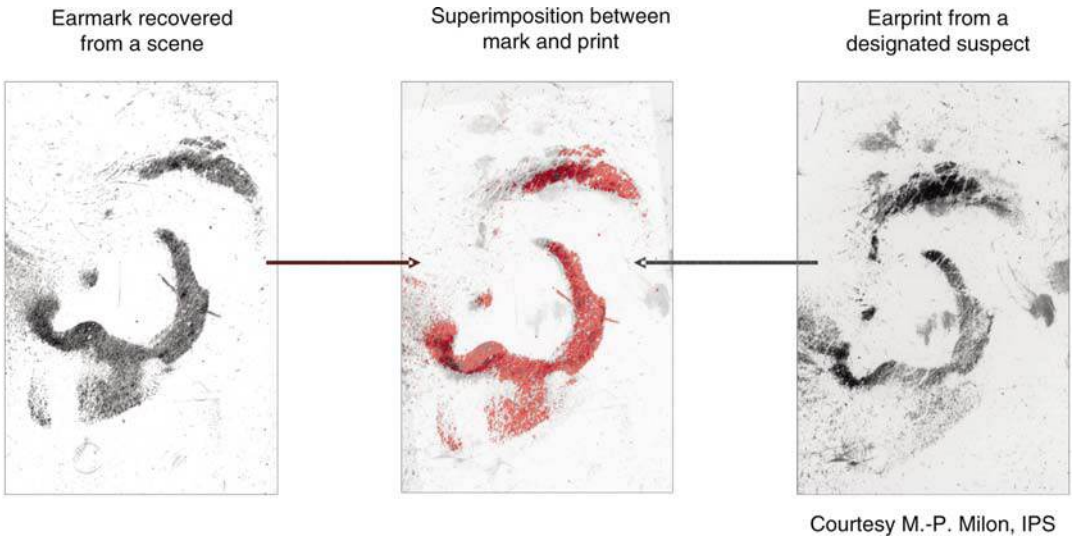
The focus here will be on this evaluative use of earprint evidence as a means to guide to the establishment of the identity of the donor of the recovered earmark(s).

Current Practice of Earmark to Earprints Comparison

The protocol used by practitioners to compare earmark(s) and earprints corresponds to the ▶ ACE-V process used in other identification fields (e.g., fingerprints) [1]. It can be summarized through the following steps:

1. The earmarks and the earprints are evaluated to assess which parts or features are visible and constitute pressure points. A specialized terminology is used to designate the anatomical parts of the ear that came into contact with the substrate (Fig. 1). Pressure points correspond to the cartilaginous parts of the ear that came into contact with the surface. The pressure on these parts is generally higher than that on the soft tissues, hence producing signs of stronger pressures on the mark as well. Also, because cartilaginous parts are less malleable than the soft tissues (such as the lobule), these pressure points tend to be more limited within source variability.
2. Because the ear is a flexible three-dimensional object, consisting of a cartilage and a covering skin, pressure of application and rotation of the head cause differences between the successive prints from the same individual. Hence, an examination of a series of known earprints from one donor, taken under various conditions, allows the creation of an empirical model of the expected variations caused by pressure and distortion (Fig. 2). This analysis will set the tolerances that will be applied during the comparison process either to retain a potential donor as a "match" or to exclude him or her as being the contributor.
3. The earmark is compared with the earprints using overlays. The examiners look at the agreement in pressure points and measurements. The more stable features being: the crus of the helix; the tragus; and the anti-tragus. They act as anchoring points for the overlay.
4. Differences in the comparison process are evaluated by the examiners in the light of the tolerances defined by the known effect of pressure and distortion. A decision is made as to whether any difference is significant (hence leading to an exclusion) or can be accounted for (hence leading to a

Earprints, Forensic Evidence of. **Figure 2** Earprints taken from a given individual with three degrees of pressures.



Earprints, Forensic Evidence of. **Figure 3** Demonstration of the correspondence between an earmark and an earprint using the superimposition technique.

"match"). The assessment of potential differences between marks and prints is left to the examiner's judgement.

5. From the quality and extensiveness of the overlay, a judgement is made as to whether the earmark and the earprints share a common origin.
6. The demonstration of the association is provided either by transparency overlays and using montages made of cut out photographs (mark and print) or using video overlays (Fig. 3).

The identification process is described mainly as a matching process – an assessment of the adequacy of superimposition between the mark and the prints – but the crucial question of the value to be given to a match is left to the examiner's judgement. In other words, when a match is declared, the assessment of the rarity of the shared features taking into account the tolerances relies on the examiner's experience.

Critical Analysis

Earmark to earprints comparison relies at the moment more on individual experience and judgement than on a structured body of research undertaken following strict scientific guidelines. The recognition process is highly subjective that exploits the extraordinary power of the human eye-brain combination.

Compared to established identification fields, such as fingerprints or handwriting comparison, the body of literature pertaining to earmarks identification is rather limited. About 60 papers have been published, a limited number in recent peer-reviewed journals. Scientific research has been mainly devoted to the study of the variability of ear morphology based on the examination of photographs of ear. The relevance of this body of knowledge to cases involving ear impressions found for example on window panes is rather limited.

Most published studies on earprints have been carried out on photographs of ears and not on the earprints or earmarks [1, 2]. The limitations of such studies are obvious there is an attempt to apply these data to the assessment of earmark to earprint comparisons for the following reasons:

- Numerous morphological features of the ear are not discernible (or cannot be classified) on earmarks.
- It is not feasible to carry out many measurements on earmarks.
- The within-source variability of features and measurements has not been fully investigated (variability, observed on marks of the same person, caused by the process of leaving and recovering marks).
- The same applies to the assessment between-persons variability (how marks from different donors can be distinguished). It is expected that

the distinguishability of earmarks from different persons will be much lower than what is observed on photographs of the ear.

There is no vast empirical study exploring the chance of finding indistinguishable marks left by different individual ears. The field of earmark identification is at its infancy and would benefit from a structured program of research.

Admissibility in Court

Within the European community, there is no specific admissibility rule regarding scientific evidence (in contrast to the *Frye*/*Daubert* standard in the United States of America [3]). The principle of the judges' free evaluation of the evidence prevails. Hence, it is not surprising to see limited debate in the European jurisprudence regarding the admissibility of the earprint evidence. The current casuistic in Switzerland (known historically for the use of earprints in criminal investigations [4, 5]) gives a contrasted view between the cases where earprint has been used in court for identification (Geneva) and where the prosecution refrained from using the evidence because of its limited contribution to address the issue of identity of sources (Ticino). Earprint evidence has also been used and accepted in the courtrooms of Belgium and the Netherlands.

In the United Kindgom, two cases involving earmarks have reached the Court of Appeal. The Court of Appeal in *R. v. Dallagher* [6] allowed the admission of earprint evidence but received additional information that emerged more clearly since the first trial that shed some new light on the strength of the evidence. Had that evidence been available to the defence at trial, it might have reasonably affected the decision of the jury to convict and hence the conviction was quashed and a new trial was ordered [7]. The Court however ruled that earprint evidence was held admissible, leaving the duty of highlighting its limits to the adversarial system itself through a proper *voir dire* or at trial. That decision was confirmed in *R. v. Mark J. Kempster* [8].

In the American case *State vs. Kunze* [9] the Court heard some twenty experts in identification evidence and came to the conclusion that earmark identification was not a field that has gained general acceptance among peers. The Court ruled that earmark evidence cannot be accepted as scientific evidence under the *Frye* test. The re-investigation of this case led to the discovery of close neighbours (close agreement between earmarks originating from different sources) among the potential donors in that case [10].

Recent Research Initiatives

Early efforts towards systematic classification or matching of earprints focused on an extraction of shape features in the anthelix area and a concept of a database based on 800 earprints from different individuals.

The field of earprint identification has been recently researched through an important initiative under funding of the European Community PF6 programme FearID (http://artform.hud.ac.uk/projects/fearid/fearid.htm?PHPSESSID = 9c4fd025eec23ee10262d9e226ff73d0).

They showed encouraging discriminative power, but without fully addressing the issue of within donor variation. Meijerman et al. showed the extent of changes on earprints features in terms of size and position [11]. The main source of intra-individual variation in earprints is the variation in pressure that is applied by the ear to the surface during listening. Studies in applied force while listening showed that intra-individual variation in applied force is comparatively small compared with the inter-individual variation [12, 13].

Semi-automatic acquisition of earprint features was also undertaken within the FearID research programme. The definition of the feature vector relied on the annotation of earprint images by skilled operators. Between-operator variations were causing a large detrimental effect on the efficiency of the system [14]. The efficiency of the developed recognition system has been tested [15]. The features are extracted from a "polyline" superimposed on the earprint by an operator. The matching is obtained using Vector Template Machine (described in http://forensic.to/fearid/VTMfinal.doc). For print to print comparisons, it was shown that for 90% of all query searches the best hit is in the top 0.1% of the list. The results become less favorable (equal error rate of 9%) for mark to print comparisons.

In addition to the described semi-automated approaches, fully automatic methods have been initially tested on a limited sample of 36 right earprints from six pairs of identical twins [16] using *Keypoint Matching* algorithms.

Some landmark research in ear biometrics [17–21] is also expected to have a drastic impact on the forensic research in earmarks in the years to come.

Related Entries

► Ear Biometrics, 3D
► Physical Analogies for Ear Recognition

References

1. van der Lugt, C.: Earprint Identification. Elsevier Bedrijfsinformatie, Gravenhage (2001)
2. Iannarelli, A.V.: Ear Identification. Paramont Publishing Company, Fremont, CA (1989)
3. Berger, M.A.: The Supreme Court's Trilogy on the Admissibility of Expert Testimony. In: Federal Judicial Center (ed.) Reference Manual on Scientific Evidence. Federal Judicial Center, Washington, 9–38 (2000)
4. Hirschi, F.: Identifizierung von Ohrenabdrüken. Kriminalistik. **24**(2), 75–79 (1970)
5. Hirschi, F.: Cambrioleurs internationaux convaincus à l'aide de preuves peu communes. Revue internationale de police criminelle **25**(239), 184–193 (1970)
6. R. *v.* Mark Dallagher, UK Court of Appeal, EWCA Crim 1903, July 25
7. Anon. Cases in Brief. Archbold News. 2003; September 19(8)
8. R. *v.* Mark J. Kempster, UK Court of Appeal, EWCA Crim 3555
9. State *v.* D. W. Kunze, Court of Appeals of Washington, Division 2, 97 Wash. App. 832, 988 P.2d 977
10. Cwiklik, C., Sweeney, K.M.: Ear Print Evidence: State of Washington v. Kunze. Personal communication from Cwiklik & Associates, 2400 Sixth Avenue South #257, Seattle, WA 98134; 2003
11. Meijerman, L., Sholl, S., De Conti, F., Giacon, M., van der Lugt, C., Drusini, A., et al.: Exploratory Study on Classification and Individualisation of Earprints. Forensic Sci Int. **40**, 91–99 (2004)
12. Meijerman, L., Nagelkerke, N., Brand, R., van der Lugt, C., van Basten, R., De Conti, F.: Exploring the Effect of Occurrence of Sound on Force Applied by the Ear when Listening at a Surface. Forensic Science, Medicine and Pathology. **1**(3), 187–192 (2005)
13. Meijerman, L., Nagelkerke, N., van Basten, R., van der Lugt, C., De Conti, F., Drusini, A., et al.: Inter- and Intra-Individual Variation in Applied Force when Listening at a Surface, and Resulting Variation in Earprints. Medicine Science and the Law. **46**(2), 141–151 (2006)
14. Alberink, I.B., Ruifrok, A.C.C., Kieckhoefer, H.: Interoperator Test for Anatomical Annotation of Earprints. Journal of Forensic Sciences. **51**(6), 1246–1254 (2006)
15. Alberink, I., Ruifrok, A.: Performance of the FearID earprint identification system. Forensic Science International. **166**(2–3), 145–154 (2007)
16. Meijerman, L., Thean, A., van der Lugt, C., van Munster, R., van Antwerpen, G., Maat, G.: Individualization of earprints. Forensic Science, Medicine, and Pathology. **2**(1), 39–49 (2006)
17. Choras, M.: Ear Biometrics Based on Geometrical Method of Feature Extraction. AMDO. 2004;LNCS 3179:51–61
18. Pun, K.H., Moon, Y.S.: Recent Advances in Ear Biometrics. In: Proceeding of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FRG'04). 2004
19. Choras, M.: Ear Biometrics Based on Geometrical Feature Extraction. Electronic Letters on Computer Vision and Image Analysis. **5**(3), 84–95 (2005)
20. Hurley, D.J., Nixon, M.S., Carter, J.N.P: Force Field Feature Extraction for Ear Biometrics. Computer Vision and Image Understanding. **98**, 491–512 (2005)
21. Lammi, H-K.: Ear Biometrics. Lappeenranta, Finland: Lappeenranta University of Technology (2005)

e-Authentication, Remote Access (Partial)

► Remote Authentication

Eigenface

Eigenface is a digitized set of face templates. The images are at the same pixel resolution and taken under standardized lighting levels and scaled to align the eyes and mouth. Any human face can be considered to be a combination of these standardized face templates. Storage capacity can be greatly improved as faces can be recorded as a list of values pertaining to

the percentage value that each eigenface contributes towards the target face.

► Face, Forensic Evidence of
► Face Sample Quality

Elastically Adaptive Deformable Model

Deformable models are 2D or 3D models that offer a data-driven recovery process in which forces deform the model until it fits the data. Global deformation parameters represent the salient shape features of natural parts, and local deformation parameters capture shape details. Instead of having the user determine the deformation parameters, in an elastically adaptive deformable model the elastic parameter values of the deformable models are determined automatically. In particular, the elastic parameters decrease when the model does not fit the data, and increase when the model is close to the data.

► Face Recognition, 3D-Based

Electromagnetic Radiation

Another term for light; fluctuations of electric and magnetic fields in space.

► Face Recognition, Thermal

Electromagnetic Resonance

Electromagnetic resonance is a phenomenon produced by simultaneously applying steady magnetic field and electromagnetic radiation (usually radio waves) to a sample of electrons and then adjusting both the strength of the magnetic field and the frequency of the radiation to produce absorption of the radiation. The resonance refers to the enhancement of the absorption that occurs when the correct combination of field and frequency is obtained.

► Digitizing Tablet

Electromagnetic Spectrum

The universe contains a vast (infinite) range of electromagnetic waves commonly referred to as the electromagnetic spectrum. At the low frequency end of the spectrum there are radio waves with wavelengths measured in metres or even kilometres. As the frequency increases and the wavelength decreases (frequency $f = c/\lambda$ where c = speed of light (3×10^8 m s^{-1}) and λ is the wavelength in metres) the electromagnetic waves are referred to as microwaves, infrared, visible light, ultraviolet, X-rays, and finally Gamma rays. At the high frequency end of the spectrum Gamma rays have a wavelength λ of the order 1×10^{-12} m. Visible light is only a very small range of the electromagnetic spectrum with wavelength from about 400 to 700×10^{-9} m.

► Face Recognition, Thermal
► Hand Veins

Embedded Processor

► Embedded Systems

Embedded Software

► Embedded Systems

Embedded Systems

Naohisa Komatsu[1], Manabu Nakano[2]
[1]Waseda University, Shinjuku-ku, Tokyo, Japan
[2]Information-technology Promotion Agency (IPA), Bunkyo-ku, Tokyo, Japan

Synonyms

Embedded processor; Embedded software

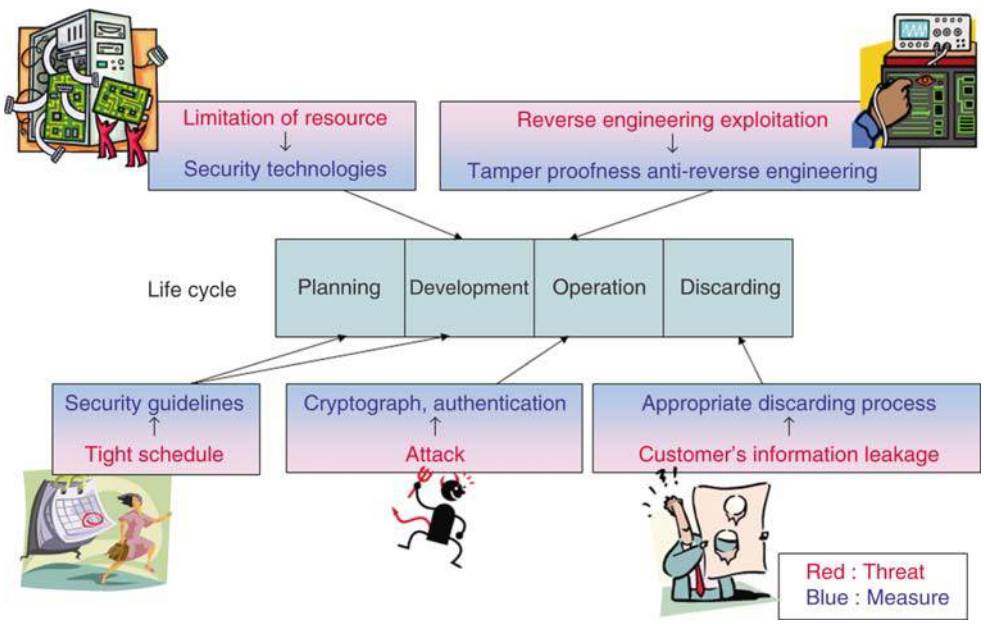
Definition

Embedded systems [1, 2] are computer systems that are embedded in various parts of equipment to control them. There is also another definition: embedded systems are integrated systems that are combined with equipment. Examples of equipment to which embedded technologies are applied include electrical household appliances and electrical equipment, PC peripheral equipment, office automation equipment, communications equipment, network facilities, medical equipment, and robots. Embedded systems are rapidly spreading wide to include social life, but there are some problems. The greatest challenge is to keep or improve the quality of design and reliability as the systems get large and complex. Biometric authentication functions have been already embedded in smart cards and cellular phones. Embedded authentication functions are applied to the driving system and personal comfort equipment at home; system security and usability are other important aspects to be studied.

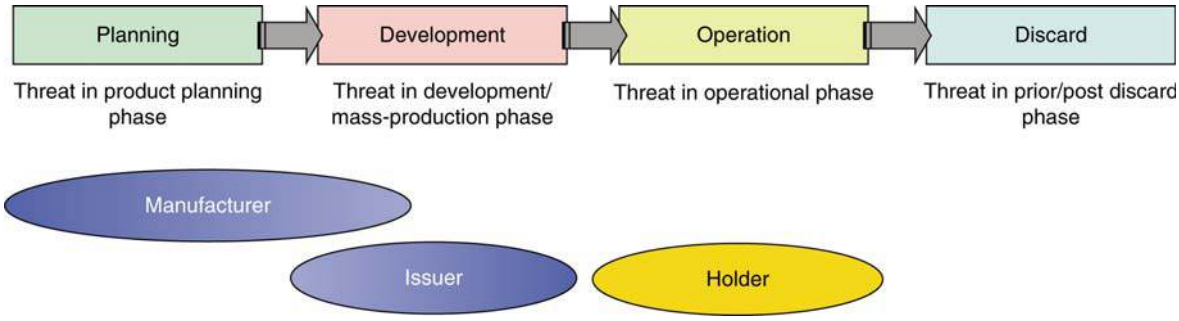
Profile of Embedded System

Those devices that are traditionally controlled by hardware-like logic have advanced significantly with the use of super micro computers and their control software since the 1980s. As a result, any complicated embedded system can be created, even in a small space and at low cost, : every device, such as home appliances, mobile phones, vehicles, industrial robots, is being popularized as an "embedded system". Biometric products are also considered "embedded systems" and will be embedded in a variety of devices such as vehicles, mobile phones, etc. in the near future. In addition, because of advanced IT technology, it is becoming easy to include communication functions; "embedded systems" are evolving as one of the infrastructural devices in ubiquitous networks, allowing us to utilize networks anytime and anywhere.

Add-on systems are defined as hardware systems in which certain software is installed, upon procurement from its manufacturer. In biometric authentication,



Embedded Systems. Figure 1 Potential ploblems & measures.



Embedded Systems. **Figure 2** Lifecycle of IC Card.

the software is referred to as the authentication software, and enables us to characterize biometric data, and cross-check biometric data between and the driver, which controls the sensor for biometric authentication. As the devices integrated with such software are commonly employed in biometric authentication products, most often, biometric authentication is done by add-on systems.

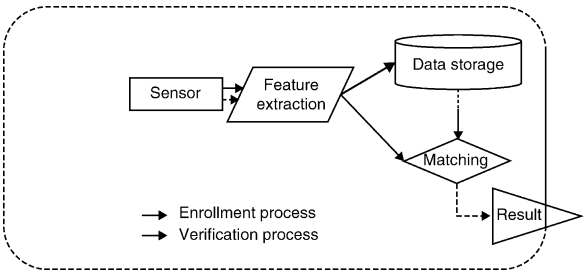
The product/system for biometric authentication can generally be classified into the following four categories, depending on how the biometric information sensor and the software that can serve authentication and/or the biometric data memory, which stores individual biometric data, are integrated with the entity (i.e., the system) that implements their original objectives upon procurement.

All-in-One

All the three – the biometric information sensor, the biometric authentication software, and the biometric data memory – are integrated with the entity (i.e., the system). The stand-alone laptop type computer with finger print authentication sensor, a door security device, and a car with finger print authentication are in the family of such biometric authentication products. The mobile telephone with finger print authentication, which is popular in Japan, is also an all-in-one biometric authentication product.

Biometric-Information-Data-Separation Method

The biometric information sensor and the biometric authentication software are integrated with the entity,



Embedded Systems. **Figure 3** Vulnerabilities in Biometric Systems.

but the biometric data memory is located separately. The biometric memory may be a handheld type of memory medium such as a smartcard and/or the server in a server–client system. The method of using a smartcard as a biometric data memory is referred to as STOC (store on card) authentication method.

Authentication-Sequestration Method

In this method, only the biometric authentication sensor is integrated with the entity (i.e., the system), but the biometric authentication software and the biometric data memory are located separately. That is, the biometric data fed by the biometric authentication sensor is transferred to a different system/device where biometric data are registered and cross-checked. As for the different system/device, a smartcard and/or the server, and part of a server–client system are included. As the smartcard itself is a device, the authentication software and/or individual data memory with the authentication method is exclusively referred to as MOC (match on card) authentication method.

Authentication-Unit

The biometric authentication sensor, biometric authentication software, and biometric data memory form a unit. It may be configured after providing the SIer and third party (ies) biometric authentication mechanism. This can provide vendors and SIer with the most simple and manageable biometric authentication.

Embedded Systems. **Table 1** Biometric-specific Vulnerabilities

Name of Vulnerability	Definitions
Familiarity/Proficiency	Certain familiarity/proficiency is necessary upon utilizing biometric system
Acceptability	Some users are still reluctant to use biometric system
FAR (False Accept Rate)	Accidental occurrence of FAR
FRR (False Reject Rate)	Accidental occurrence of FRR
Unavailability	There are some users who cannot be authenticated biologically or are those for which biometric data cannot be obtainable from
User Status	Data granularity will vary depending on user's physical status
Entering Environment (Minutia Angle, etc.)	Data granularity will vary depending on entering environment, such as minutia angle, etc.
Wolf	FAR occurs with high probability due to Wolf
Lamb	FAR occurs with high probability due to Lamb
Goat	FRR occurs with high probability due to Goat
Authentication Parameter	Inadequate matching performance relevant to configuration of authentication parameter
Falsified-biometric Information	Physically generation of falsified biometric information
Publication	Anyone else can acquire user's biometric information
Assumption	Assumable biometric information from templates/matching results
Extent	Number of attempts available to biometric information/user/authentication
Similarity	There are some users whose biometric information is nearly identical to others

Embedded Systems. **Table 2** Vulnerabilities Common to General IT Systems

Name of Vulnerability	Definitions
Registration	Vulnerability upon registration
Singularity	Available to attack against anyone else's IDs without any tools when biometric information is simply used
Alternative Means	There always need certain means alternable for biometric authentication as there are some people who cannot be authenticated by or there are those whose biometric data cannot be obtainable from
Presence	Biometric information is presentable to third party/people if the owner grants
Motivation	Verifiable/identifiable data entry is necessary by the user of biometric system
Sensor Exposure	Sensor which collecting biometric data is disclosed to outside
Data Leakage	Leakage of biometric data stored in biometric system to outside
Side-channel	Leakage of the information relevant to biometric system to outside
Data Alteration	Alteration availability for those data stored in biometric system
Configuration Management	Upon differed conformity in elements which configuring system, normal operation and matching performance required are getting disabled
Deactivation	Authentication is getting unavailable temporarily when some parameters are satisfied

Difference Between Embedded System and General System

What if we do not conduct certain security measures on the general computer systems connected to a network? They will be infected with virus within a short period of time and/or they will be easily attacked by malicious persons. To avoid that, it is necessary to conduct certain security measures, utilizing antivirus software, firewall, etc., and in case of some vulnerability in software, we can maintain security by downloading and applying security patches in general systems. However, in an "embedded system", it is harder to address the said measures because of the constraints in utilizing their resources. In addition, there are the "embedded system"-specific issues such as side-channel attack and ▶ reverse engineering.

It is expected that along with advancement, such security issues looming up in the world of computer systems will be a great threat to the "embedded system" in the years to come. There are only some accidents relevant to the "embedded system" that have been reported and it is not likely that they will happenfrequently hereafter. It is ideal to construct the lifecycle of the "embedded system" in four different phases: planning, development, operation, and discarding, to implement sufficient security measures by both developers and users (Fig. 1).

Embedded Systems. **Table 3** Lifecycle of Embedded Systems and their Vulnerabilities

	Name of Vulnerability	Planning	Development	Operational	Discard
Biometric System-Specific Vulnerabilities	Familiarity/Proficiency			Y	
	Acceptability			Y	
	FAR (False Acceptance Rate)			Y	
	FRR (False Resistance Rate)			Y	
	Unavailability			Y	
	User Status			Y	
	Entering Environment (Minutia Angle, etc.)			Y	
	Wolf			Y	
	Lamb			Y	
	Goat			Y	
	Authentication Parameter			Y	
	Falsified-Biometric Information			Y	
	Publication			Y	
	Assumption			Y	Y
	Extent	Y	Y		
	Similarity			Y	
Vulnerabilities Common to General IT Systems	Registration			Y	
	Singularity			Y	
	Alternative Means			Y	
	Presence			Y	Y
	Motivation			Y	
	Sensor Exposure		Y	Y	
	Data Leakage			Y	Y
	Side-channel			Y	Y
	Data Alteration		Y	Y	
	Configuration Management		Y	Y	
	Deactivation			Y	

Instances of Embedded System in IC Card

In this section, the essay introduces the IC card as one of the instances of specific embedded systems. Generally, it can be assumed that the "manufacturer", "issuer," and "holder" will be involved in the lifecycle of the IC card from its planning to discarding phases (Fig. 2).

There exist different threats in each phase of the lifecycle:

– Planning Phase:	The leakage of Information relevant to design document
– Development Phase:	Fraudulent issuance of public key certificate
– Operational Phase:	Compromise in cryptography protocol for long service
– Discard Phase:	The deprivation of private information stored in the card

To avoid these problems, it is necessary to adopt certain security measures such as encryption of the design document, periodic logical verification, and regulation of prior/post disposal, etc. for the respective holders' further security. For effective security, all the measures have to be employed to work synergically.

Instances in Biometrics

Figure 3 shows the vulnerabilities of a biometric authentication system [3, 4], and the vulnerabilities [5] are explained in Tables 1 and 2. Vulnerabilities can be classified into two types: those biometric-specific and those common to general information systems. However, in the latter case, only those that may cause a threat when combined with the biometric-specific vulnerability(ies) are listed. Table 3 shows the vulnerabilities in biometric systems in the respective phases: Planning, Development, Operational, and Discard.

Related Entries

► Biometrics, Overview

References

1. Henzinger, T.A., Sifakis, J.: The discipline of embedded systems design, IEEE Comput. pp. 32–40 (Oct. 2007)
2. Hwang, D.D., Schaumont, P., Tiri, K., Verbauwhede, I.: Securing embedded systems, IEEE Security & Privacy, pp. 40-49 (Mar./Apr. 2006)
3. Jain, A.K., Bolle, R., Pankanti, S.: Biometrics: Personal Identification in Networked Society, Kluwer Academic Publishers (1999)
4. Jain, A.K., Ross, A., Pankanti, S.: An introduction to biometric recognition. IEEE T. Circ. Syst. Vid. **14**(1), pp. 4–20 (2004)
5. Faundez-Zanuy, M.: On the vulnerability of biometric security systems, IEEE Aero. El. Sys. Mag. **19**(6), pp. 3–8 (2004)

Embedding Space

Embedding space is the space in which the data is embedded after dimensionality reduction. Its dimensionality is typically lower that than of the ambient space.

► Manifold Learning

Empirical Analysis

Empirical analysis in the context of biometric sample synthesis deals with the creation of parametric or mathematical models, which mimic natural statistical factors such as the density of distinguishing features such as bifurcations in fingerprints, or the shapes of ears. In such cases, many models describing these biometric patterns are derived from the observational analysis of real patterns instead of on physical laws governing their creation or growth.

► Biometric Sample Synthesis

Empirical Statistical Models

Empirical statistical models attempt to recreate real world distributions based on the empirical analysis of

a population of real samples. In biometric sample synthesis an example of an empirical statistical model is the frequencies of loops and whorls on various fingers.

► Biometric Sample Synthesis

Encoded Finger Data

► Finger Data Interchange Format, Standardization

Encoder

Encoder is a software that extracts the features from a fingerprint image.

► Universal Latent Workstation

Encoding of Hand Geometry Information

► Hand Data Interchange Format, Standardization

Encryption, Biometric

Ann Cavoukian, Alex Stoianov
Office of the Information and Privacy Commissioner, Toronto, ON, Canada

Synonyms

Biometric cryptosystem; Biometric key generation; Biometric locking; Fuzzy extractor; Secure sketch

Definition

Biometric Encryption (BE) is a group of emerging technologies that securely bind a digital key to a biometric or generate a digital key from the biometric, so that no biometric image or template is stored. What is stored is the BE template otherwise known as a "biometrically encrypted key" or "helper data." As a result, neither the digital key nor the biometric can be retrieved from the stored BE template. BE conceptually differs from other systems that encrypt biometric images or templates using conventional encryption, or store a cryptographic key and release it upon successful biometric authentication. With BE, the digital key is recreated only if the correct biometric sample is presented on verification. The output of BE verification is either a digital key or a failure message. This "encryption/decryption" process is fuzzy because of the natural variability of biometric samples. Currently, any viable BE system requires that biometric-dependent helper data be stored.

Introduction

Biometric technologies may add a new level of authentication and identification to applications, but are not, however, without their risks and challenges. There are important technological challenges such as accuracy, reliability, data security, user acceptance, cost, and interoperability, as well as challenges associated with ensuring effective privacy protections. Some common security vulnerabilities of biometric systems include:

Spoofing; replay attacks; substitution attacks; tampering; masquerade attacks (creating a digital "artifact" image from a fingerprint template so that this artifact, if submitted to the system, will produce a match); Trojan horse attacks; and overriding Yes/No response (which is an inherent flaw of existing biometric systems).

In addition to the security threats that undermine the reliability of biometric systems, there are a number of specific privacy concerns with these technologies:

- function creep (i.e., unauthorized secondary uses of biometric data)
- expanded surveillance, tracking, profiling, and potential discrimination (biometric data can be matched against samples collected and stored elsewhere and used to make decisions about individuals)

- data misuse (data breach, identity theft, and fraud)
- negative personal impacts of false matches, non-matches, system errors, and failures (the consequences of system anomalies, especially in large-scale systems, often fall disproportionately on individuals, normally in the form of inconveniences, costs, and stigma)
- insufficient oversight, accountability, and openness in biometric data systems
- potential for collection and use of biometric data without knowledge, consent, or personal control

These types of risks threaten user confidence, which leads to a lack of acceptance and trust in biometric systems.

Biometric Encryption (BE) technologies can help to overcome the prevailing "zero-sum" mentality involved in traditional biometrics, namely, that adding privacy to authentication and information systems weakens security. With BE, it is possible to enhance both privacy and security in a positive-sum model.

What is Biometric Encryption (BE)?

The concept of Biometric Encryption (BE) was first introduced in the mid-90s by G. Tomko et al. [1]. For more information on BE and related technologies, see the review papers in [2–4].

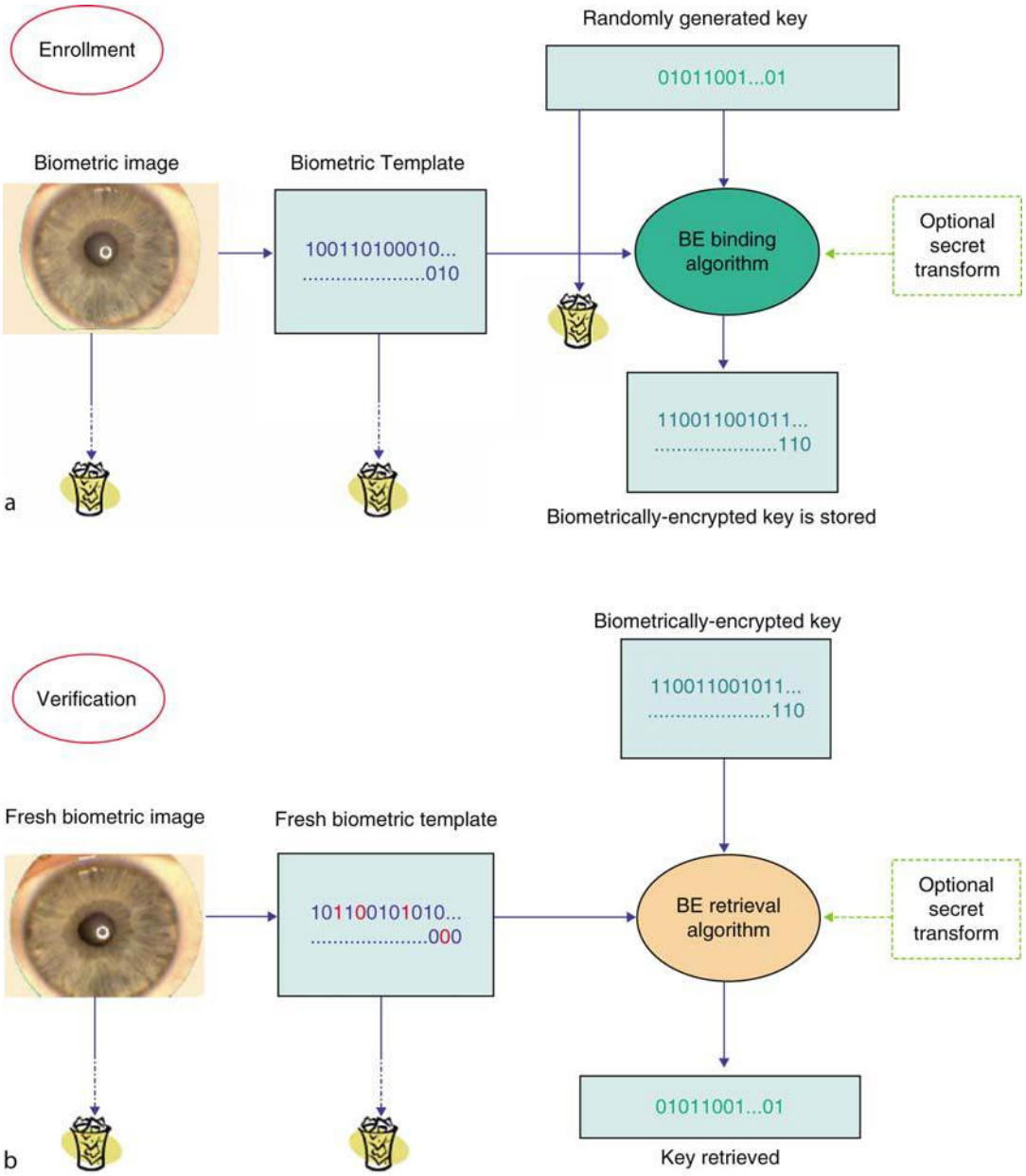
Biometric Encryption is a process that securely binds a digital key to a biometric or generates a key from the biometric. In essence, the key is "encrypted" with the biometric, and the resulting biometrically encrypted key, also called BE template or helper data, is stored. The digital key can be "decrypted" on verification if a correct biometric sample is presented. This "encryption/decryption" process is fuzzy by nature, because the biometric sample is different each time, unlike an encryption key in conventional cryptography. A major technological challenge is to have the same digital key recreated despite the natural variations in the input biometrics.

After the digital key is recreated on verification, it can be used as the basis for any physical or logical application. The most obvious use is in a conventional cryptosystem where the key serves as a password and may generate, for example, a pair of Public and Private keys. It should be noted that BE itself is not a cryptographic algorithm. The role of BE is to replace or augment vulnerable password-based schemes with more secure and more convenient biometrically managed keys.

BE should not be mistaken for other systems that encrypt biometric images or templates using conventional encryption, or store a cryptographic key in a trusted token/device and subsequently release it upon successful biometric verification (i.e., after receiving Yes response). However, BE is related to another family of privacy-enhancing technologies called ▶ Cancelable Biometrics (CB) (N. Ratha et al. in [3]; see also the Encyclopedia article on "Cancellable Biometrics"). CB applies a transform (preferably, noninvertible) to a biometric image or template and matches the CB templates in the transformed domain. This transform is usually kept secret. Unlike BE, the CB system does not bind or generate a key. CB remains inherently vulnerable to overriding Yes/No response and to a substitution attack.

There are two BE approaches: key binding, when an arbitrary key (e.g., randomly generated) is securely bound to the biometric, and key generation, when a key is derived from the biometric. Both approaches usually store biometric dependent helper data. Some BE schemes (e.g., Fuzzy Commitment [5], Fuzzy Vault [6]) can equally work in both key generation and key binding mode; the key generation is also called "secure sketch" or "fuzzy extractor" as defined in [7]. Secure sketch implies that the enrolled biometric template will be recovered on verification when a fresh biometric sample is applied to the helper data (i.e., the enrolled template itself or a string derived from it, e.g., by hashing the template, serves as a digital key). Note, however, that this "key" is not something inherent or absolute for this particular biometric; it will change upon each re-enrolment. The size of the key space for the secure sketch is defined by the intraclass variations of the biometric as opposed to the key binding approach.

In the key binding mode, as illustrated in Fig. 1, the digital key is randomly generated on enrollment so that neither the user nor anybody else knows it. The key itself is completely independent of biometrics, and therefore, can always be changed or updated. After a biometric sample is acquired, the BE algorithm securely and consistently binds the key to the biometric to create a biometrically encrypted key. The BE template provides privacy protection and can be stored either in a database or locally (smart card, token, laptop, cell



Encryption, Biometric. Figure 1 High level diagram of a Biometric Encryption process in a key binding mode. **(a)** Enrollment; **(b)** Verification.

phone, etc.). At the end of the enrollment, both the key and the biometric are discarded.

On verification, the user presents his or her fresh biometric sample, which, when applied to the legitimate BE template, will let the BE algorithm recreate the same key. At the end of verification, the biometric sample is discarded once again. The BE algorithm is designed to account for acceptable variations in the input biometric. On the other hand, an impostor whose biometric sample is different enough will not be able to recreate the key.

Many BE schemes also store a hashed value of the key (not shown in Fig. 1) so that a correct key is released from the BE system only if the hashed value obtained on verification is exactly the same. Also, good practice would be not to release the key, but rather, another hashed version of it for any application. This hashed version can in turn serve as a cryptographic

key. With this architecture, an attacker would not be able to obtain the original key outside the BE system. Likewise, the biometric image/template should not be sent to a server; the BE verification should be done locally in most scenarios.

An important part of most BE algorithms is an Error Correcting Code (ECC). ECCs are used in communications, for data storage, and in other systems where errors can occur. Biometric Encryption is a new area for the application of ECC. For example, a binary block ECC, which is denoted (n, k, d), encodes k bits with n > k bits by adding some redundancy. Those n-bit strings are called codewords; there are 2^k of them in total, where k is the key length. The minimum distance (usually a Hamming distance is implied) between the codewords is d. If, at a later stage (in case of BE, on verification), the errors occur, the ECC is guaranteed to correct up to (d-1)/2 bit errors among n bits. Ideally, the legitimate users will have a number of errors within the ECC bound so that the ECC will decode the original codeword, and hence, the digital key. On the other hand, the impostors will produce an uncorrectable number of errors, in which case the ECC (and the BE algorithm as a whole) will declare a failure. In practice, BE, like any biometric system, has both false rejection and false acceptance rates (FRR and FAR). Note that BE does not use any matching score; instead, the FRR/FAR tradeoff may be achieved in some cases by varying the parameters of the BE scheme. Some ECCs may work in a soft decoding mode, that is, the decoder always outputs the nearest codeword, even if it is beyond the ECC bound. This allows achieving better error-correcting capabilities.

To improve the security of a BE system, an optional "transform-in-the-middle" (shown in the dashed square in Fig. 1) may be applied. Preferably, the transform should be non-invertible and kept secret. One of the ways would be employing a randomization technique, such as Biohashing [8] or "salting" in more general terms [2]. The transform can be controlled with the user's password or can be separated from the rest of the helper data by storing it on a token or a server.

Advantages and Possible Applications of BE

BE technologies can enhance both privacy and security in the following ways:

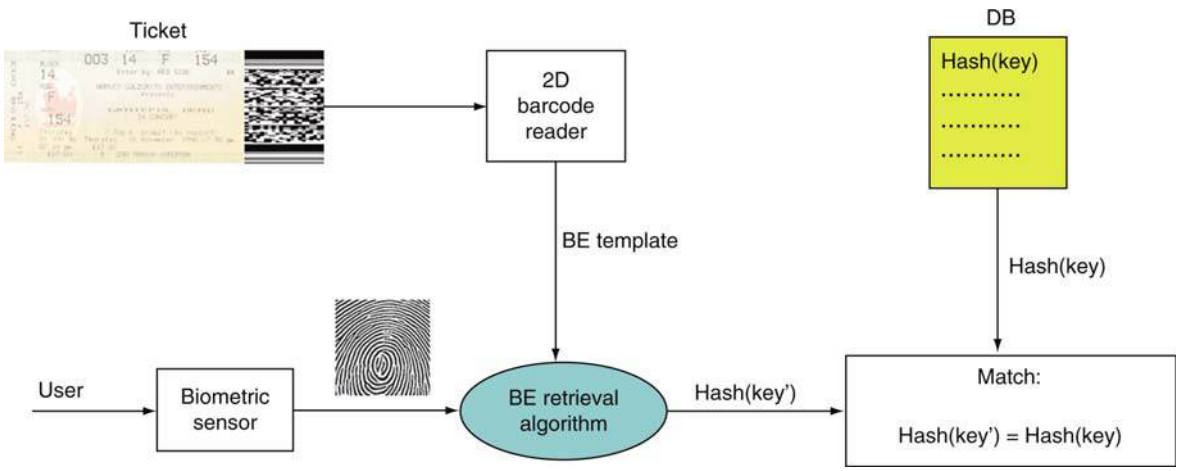
- There is no retention of biometric image or conventional biometric template, and they cannot be recreated from the stored helper data.
- They are capable of multiple identifiers: a large number of BE templates for the same biometric can be created for different applications.
- The BE templates from different applications cannot be linked.
- The BE template can be revoked or canceled.
- They can be easily integrated into conventional cryptosystems, as the passwords are replaced with longer digital keys, which do not have to be memorized.
- They provide improved authentication and personal data security through a stronger binding of user biometric and system identifier.
- The BE systems are inherently protected from substitution attack, tampering, Trojan horse attack, overriding Yes/No response, and less susceptible to masquerade attack.
- They are suitable for large-scale applications, as the databases will store only untraceable, yet sufficient, information to verify the individual's claim.

These features embody standard fair information principles, providing user control, data minimization, and data security.

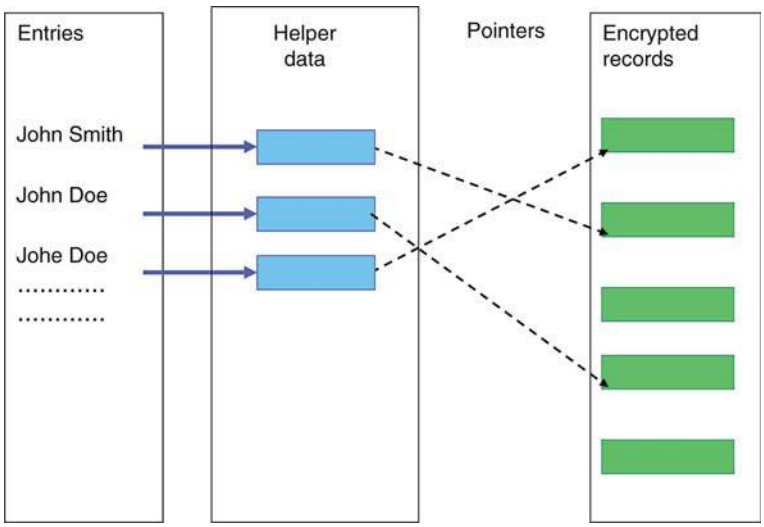
As such, BE technologies put biometric data firmly under the exclusive control of the individual, in a way that benefits the individual and minimizes the risk of function creep and identity theft. They provide a foundation for building greater public confidence, acceptance, and use, and enable greater compliance with privacy and data protection laws.

Possible applications and uses of Biometric Encryption include

- Biometric ticketing (Fig. 2) for events
- Biometric boarding cards for travel
- Drug prescriptions
- Three-way check of travel documents
- Identification, credit, and loyalty card systems
- Anonymous databases (Fig. 3), that is, anonymous (untraceable) labeling of sensitive records (medical, financial)
- Consumer biometric payment systems
- Remote authentication via challenge-response scheme
- Access control (physical and logical)



Encryption, Biometric. **Figure 2** Biometric ticketing. A BE template is stored on a ticket as a 2D bar code, and a database stores the hashed value of a key, Hash(key), for each enrolled user. The key and the ticket are used only for this particular application. On a verification terminal: (i) The user presents his ticket to the system which reads in the BE template from the bar code; (ii) The live biometric sample is taken; (iii) The system applies the biometric to the BE template to retrieve the key; (iv) Hash(key') is sent to the database where it is compared to the stored version, Hash(key).



Encryption, Biometric. **Figure 3** Anonymous database controlled by Biometric Encryption. The database contains anonymous encrypted records, e.g., medical files. The cryptographic keys and the links to the entries, which may be users' names or pseudonyms, are controlled by BE. After the user enters his pseudonym, the associated BE template (helper data) is retrieved and applied to the user's biometric. If BE successfully recovers the user's digital key, it will recreate the pointer to the anonymous record and the encryption key to decrypt the record.

- Personal encryption products (i.e., encrypting files, drives, e-mails, etc.)
- Local or remote authentication of users to access files held by government and other various organizations

BE Technologies

The following are core BE schemes. The more detailed, up-to-date overviews of BE technologies are presented in [2, 4].

Mytec1

This is the first BE scheme [1]. It was developed using optical processing, but can also be implemented digitally. The key is linked to a predefined pattern, s(x), which is a sum of several delta-functions. Using s(x) and a fingerprint, f(x), one can create a filter, $H(u) = S(u)/F(u)$, in Fourier domain (S(u) and F(u) are the Fourier transforms of s(x) and f(x)). It is difficult to obtain either S(u) or F(u) from the stored filter H(u). On verification, if a correct fingerprint, $F'(u) \approx F(u)$, is applied to the filter, it will reconstruct a correct output pattern, $s'(x) \approx s(x)$ so that the key will be regenerated from the locations of the output correlation peaks. Unfortunately, this scheme turned out to be impractical in terms of providing sufficient accuracy and security.

Mytec2

This is the first practical BE scheme [9]. Unlike Mytec1, it retains phase-only parts of S(u) and F(u) in the filter, H(u). The phase of S(u) is randomly generated, but not stored anywhere. As a result, the output pattern, c(x), is also random. The key, normally 128 bit long, is linked to c(x) via a lookup table and ECC. The filter, H(u), the lookup table, and the hashed key are stored in the helper data. The system is error tolerant and translation invariant. The published version [9] used a simple repetition ECC, which makes the system vulnerable to several attacks, such as Hill Climbing [10].

However, a closer examination of the Mytec2 scheme shows that if the randomness of H(u) and c(x) is preserved on each step of the algorithm, the scheme is a variant of so-called "permutation-based fuzzy extractor" as defined in [7]. Therefore, if a proper ECC (preferably, single block) is used instead of the repetition ECC, the system will be as secure as those types of fuzzy extractors.

(Note that Mytec1 and Mytec2 schemes were originally called "Biometric Encryption", which was a trademark of Toronto-based Mytec Technologies Inc., now Bioscrypt, a fully-owned subsidiary of L1 Identity Solutions Inc. The trademark was abandoned in 2005.)

ECC Check Bits

This scheme, which was originally called "private template," is a secure sketch (i.e., a key generation) [11]. A biometric template itself serves as a cryptographic key. To account for the template variations between different biometric samples, an (n, k, d) error correcting code is used. A number of (n-k) bits, called *check bits*, are appended to the template to map the k-bit template to an n-bit codeword. The check bits are stored into the helper data along with the hashed value of the template. The scheme is impractical, since it is required that $n < 2k$ from the security perspective. Such ECC would not be powerful enough to correct a realistic number of errors for most biometrics, including iris scan.

Biometrically Hardened Passwords

This technique was developed for keystroke dynamics or voice recognition [12]. A password that the user types or says is fused with a key (via a secret sharing scheme) extracted from a biometric component, thus hardening the password with the biometrics. The technique was made adaptive by updating a "history file" (which is, in fact, helper data) upon each successful authentication. However, the types of biometrics used did not allow for achieving good accuracy numbers.

Fuzzy Commitment

This is conceptually the simplest, yet the most studied, BE scheme [5] A. Juels in [3]. A biometric template must be in the form of an ordered bit string of a fixed length. A key is mapped to an (n, k, d) ECC codeword of the same length, n, as the biometric template. The codeword and the template are XOR-ed, and the resulting n-bit string is stored into helper data along with the hashed value of the key. On verification, a fresh biometric template is XOR-ed with the stored string, and the result is decoded by the ECC. If the codeword obtained coincides with the enrolled one (this is checked by comparing the hashed values), the k-bit key is released. If not, a failure is declared.

In a "secure sketch" (i.e., key generation) mode [7], the enrolled template is recovered from the helper data on verification, if a correct (yet different) biometric sample is presented.

The scheme seems to be one of the best for the biometrics where the proper alignment of images is possible, such as iris scan [13, 14] and face recognition (T. Kevenaar in [3]). For iris, the reported results are

FRR = 0.47% at FAR $< 10^{-5}$ for a 140-bit key mapped to 2048-bit codeword [13], and FRR = 5.6% at FAR $< 10^{-5}$ (42-bit key) [14] for a poorer quality, yet more realistic, iris database.

ECC Syndrome

In this spinoff of the Fuzzy Commitment scheme, a so-called ECC syndrome of (n-k) size is stored in the helper data [7, 2]. On verification, the enrolled template is recovered (i.e., the scheme works in the secure sketch mode).

Quantization using Correction Vector

This method, which was also called "shielding functions", is applied to continuously distributed and aligned biometric features (J.-P. Linnartz et al. in [3]). For each feature, a residual is calculated, which is the distance to the center of the nearest even-odd or odd-even interval, depending on the parity of the key bit. The correction vector comprising all the residuals is stored into the helper data. On verification, a noisy feature is added to the residual and is decoded as 1 or 0, if the resulting interval is odd-even or vice versa. The scheme can work with or without (if a noise level is low) a subsequent ECC. In general, storing a correction vector could make the scheme vulnerable to score-based attacks.

Fuzzy Vault

This is, probably, the only BE scheme that is fully suitable for unordered data with arbitrary dimensionality, such as fingerprint minutiae [6, 15]. A secret message (i.e. a key) is represented as coefficients of a polynomial in a Galois field, for example, GF(2^{16}). In the most advanced version [15], the 16-bit x-coordinate value of the polynomial comprises the minutia locations and the angle, and the corresponding y-coordinates are computed as the values of the polynomial on each x. Both x and y numbers are stored alongside with chaff points that are added to hide real minutiae. On verification, a number of minutiae may coincide with some of the genuine stored points. If this number is sufficient, the full polynomial can be reconstructed using an ECC (e.g., Reed-Solomon ECC) or Lagrange interpolation. The polynomial reconstruction means that the secret has been successfully decrypted. The scheme works both in the key binding and the key generation (secure sketch) mode. The version of [15] also stores fingerprint alignment information. The best results for fingerprints show FRR = 6% – 17% at FAR = 0.02%.

The more secure version of Fuzzy Vault [7] stores high degree polynomial instead of real minutiae or chaff points. However, there are difficulties in the practical implementation of this version.

Unlike other BE schemes, the fuzzy vault actually stores real minutiae, even though they are buried inside the chaff points. This could become a source of potential vulnerabilities [2, 4]. The system security can be improved by applying a secret minutiae permutation controlled by a user's password [2]. This "transform-in-the-middle" approach is applicable to most BE schemes.

Biohashing (with key binding)

An ordered biometric feature set is transformed into a new space of a lower dimension by generating a random set of orthogonal vectors and obtaining an inner product between each vector and the biometric feature set [8]. The result (called "Biohash") is binarized to produce a bit string. The random feature vectors are generated from a random seed that is kept secret, for example, by storing it in a token. The key is bound to the Biohash via Shamir secret sharing with linear interpolation, or by using a standard Fuzzy Commitment scheme. Very good FRR/FAR numbers [8] were obtained, however, in an unrealistic "non-stolen token" scenario. Biohashing is referred more often as a CB scheme where Biohashes are matched directly, that is, without the key binding.

Graph-based Coding

In this generalization of the ECC syndrome scheme, Low Density Parity Check (LDPC) ECCs are used in a graphical representation [16]. LDPC codes, which are the state-of-the-art channel ECCs (n, k, d), can be designed with large numbers of n and k, and can handle high error rates. This makes them suitable for

BE applications. The scheme can be applied to both ordered (e.g., iris) and unordered (e.g., fingerprint minutiae) feature sets. For the latter, a factor graph models the minutiae variability as a movement, an erasure, or an insertion (i.e., spurious generation) of minutiae. The scheme uses a Belief Propagation decoding algorithm and shows promising results.

Attacks on BE

Despite the fact that many BE schemes have a formal proof of security, they may be vulnerable to low level attacks, such as when an attacker has access to helper data, is familiar with the BE algorithm, and can run the attack offline. By cracking a BE system, the attacker can pursue one or more of the following:

- Obtain the key bound to the biometrics
- Obtain the exact biometric template used on enrollment
- Obtain an approximate version of the template that, nonetheless, would defeat the system (masquerade template)
- Create a masquerade image of the biometrics
- Link BE templates generated from the same biometrics but stored in different databases

The known attacks on BE, as described in [4], are listed in the following paragraphs. Note that CB may also be vulnerable to most of the attacks.

False Acceptance attack. This is one of the "brute force" attacks. Offline, the attacker runs an impostor database of about FAR^{-1} biometric images or templates against the helper data to obtain a false acceptance. The database can be either real or computer-generated, such as *SFinGe*. The image that has generated the false acceptance will serve as a masquerade image.

Reversing the hash. This is another "brute force" attack. If a hashed key is stored into the helper data, the attacker may try to cryptographically reverse the hash. This attack should always be made more computationally expensive for an attacker than other attacks.

Hill Climbing attack [10]. Based on the knowledge of the algorithm, the attacker derives an intermediate matching score during the verification process, even though the BE algorithm does not use any score. By making small changes in the input impostor's image or template, the attacker retains the change, if the score increases, or rejects it, if not. After a number of iterations, the attacker may be able to retrieve a key and create a masquerade image/template.

The BE schemes that divide helper data into short chunks of ECC (e.g., a repetition ECC), and the schemes with a correction vector may be especially vulnerable to this and to the Nearest Impostors attack.

Nearest Impostors attack [4]. This is another score-based attack. The attacker derives a partial matching score for each ECC chunk (if any) of the helper data and a global intermediate score (like in the Hill Climbing attack). By running a small impostor database against the helper data, the attacker identifies several "nearest impostors", that is, the attempts with the highest global score, or alternatively, with the highest partial score for a given chunk. By applying a voting technique to the nearest impostors, the attacker retrieves the key bits associated with the chunk. If successful, the attack yields the entire key or at least reduces the search space for the key.

Using statistics of ECC output [4]. A small impostor database (with various distortions, rotations, and shifts applied) is run against the ECC chunks of the helper data. The number of appearances of each possible output codeword for all impostor attempts is counted to create a histogram. The codeword corresponding to the histogram maximum is declared a winner.

Using an information leak from helper data. This group of attacks may directly exploit

- Nonrandomness of the helper data [4] (e.g., if clusters in the helper data are identified, the attacker may interconnect the same parity bits)
- Alignment information and minutiae angles in the Fuzzy Vault
- A method for generating the chaff points [17]
- Nonuniformity of the output bits distribution in quantization schemes, etc.

Re-usability attack (X. Boyen in [3]). If the same biometric is re-used for different applications and/or keys, the attacker may combine several versions of the helper data to retrieve both the biometric and all the keys. Fuzzy Vault is especially vulnerable to this attack.

Among all BE schemes, it seems that one of the most secure would be a Fuzzy Commitment (or other related fuzzy extractors, such as ECC syndrome) scheme with a single block (n, k, d) ECC, where n and k are large (e.g., $n > \sim 1000$, $k > \sim 100$). From the security perspective, the amount of any additional

side information that is stored (e.g., alignment data) should be kept to a minimum.

The resilience to some of the attacks may be improved by employing the "transform-in-the-middle" approach, especially if the transform is controlled by a password/token.

Current State of BE

Many different approaches have been developed for BE, but currently few systems have been deployed or implemented into products. Until now, little work has been done to analyze the security of BE systems.

The authors' consider the following technologies as the state of the art of BE:

- Philips (the Netherlands) priv-ID™ for the face recognition (2D and 3D) and fingerprints (T. Kevenaar in [3])
- Hao et al for iris [13]
- Nandakumar et al (fuzzy vault for fingerprints) [15]
- Draper et al. of Mitsubishi Electric Research Laboratories (U.S.) for iris and fingerprints [16]
- Bringer et al of Sagem Sécurité (France) for iris [14]
- Genkey (Norway) BioCryptic® for fingerprints (unfortunately, not much information about the technology is available)

The Philips priv-ID™ technology is ready for deployment. It is part of the EU 3D Face project and of the 3-year EU TURBINE project [18]. The latter has been given significant funding and aims at piloting a fingerprint-based BE technology at an airport in Greece.

The Genkey BioCryptic® technology has been deployed for a Rickshaw project in New Delhi (India). Both Philips and Genkey systems can fit the helper data into a 2D bar code.

BE Challenges

Technologically, BE is much more challenging than conventional biometrics, since most BE schemes work in a "blind" mode (the enrolled image or template are not seen on verification). As BE advances to the next phase of creating and testing a prototype, the following issues need to be addressed:

Biometric modalities that satisfy the requirements of high entropy, low variability, possibility of alignment, and public acceptance should be chosen. At present, the most promising biometric for BE is iris followed by fingerprints and face.

The image acquisition process (the requirements are tougher for BE than for conventional biometrics) must be improved.

BE must be made resilient against attacks.

The overall accuracy and security of BE algorithms must be improved. Advances in the algorithm development in conventional biometrics and in ECCs should be applied to BE.

Multimodal approaches should be exploited.

BE applications should be developed.

Summary

Biometric Encryption is a fruitful area for research and is becoming sufficiently mature for prototype development and the consideration of applications.

BE technologies exemplify the fundamental privacy and data protection principles that are endorsed around the world, such as data minimization, user empowerment, and security.

Although introducing biometrics into information systems may result in considerable benefits, it can also introduce many new security and privacy vulnerabilities, risks, and concerns. Novel Biometric Encryption techniques can overcome many of those risks and vulnerabilities, resulting in a win-win, positive-sum model that presents distinct advantages to both security and privacy.

Related Entries

- ▶ Biometric Security, Overview
- ▶ Biometric Vulnerabilities
- ▶ Cancelable Biometrics
- ▶ SFinGe
- ▶ Template Security

References

1. Tomko, G.J., Soutar, C., Schmidt, G.J.: Fingerprint controlled public key cryptographic system. U.S. Patent 5541994, July 30, 1996 (Filing date: Sept. 7, 1994)
2. Jain, A.K., Nandakumar, K., Nagar, A.: Biometric Template Security. EURASIP J. Adv. Signal Process. v. 2008, Article ID 579416, pp. 1–17 (2008)
3. Tuyls, P., Škorić, B., Kevenaar, T. (eds.): Security with Noisy Data: Private Biometrics, Secure Key Storage and Anti-Counterfeiting. Springer, London (2007)

4. Cavoukian, A., Stoianov, A.: Biometric Encryption: The New Breed of Untraceable Biometrics. In: Boulgouris, N.V., Plataniotis, K.N., Micheli-Tzanakou, E. (eds.): Biometrics: fundamentals, theory, and systems. Wiley, London (2009)
5. Juels, A., Wattenberg, M.: A fuzzy commitment scheme. In: Tsudik G. (ed.) Sixth ACM Conference on Computer and Communications Security, pp. 28–36. ACM Press, New York (1999)
6. Juels, A., Sudan, M.: A fuzzy vault scheme. In: Lapidoth, A., Teletar, E. (eds.) Proceedings of IEEE International Symposium on Information Theory, p. 408. IEEE, Lausanne (2002)
7. Dodis, Y., Reyzin, L., Smith, A.: Fuzzy Extractors: How to Generate Strong Keys from Biometrics and other Noisy Data. In Cachin, C., Camenish, J., Proc. Eurocrypt 2004, pp. 523–540 Springer-Verlag, NY (2004)
8. Teoh, A.B.J., Ngo, D.C.L., Goh, A.: Personalised cryptographic key generation based on FaceHashing. Comput. Secur. **23**, 606–614 (2004)
9. Soutar, C., Roberge, D., Stoianov, A., Gilroy, R., Vijaya Kumar B.V.K.: Biometric Encryption (Chapter 22). In: Nichols, R.K. (ed.): ICSA Guide to Cryptography, McGraw-Hill New York, (1999)
10. Adler, A.: Vulnerabilities in Biometric Encryption Systems. In: Audio- and video-based Biometric Person Authentication (AVBPA2005). Lecture Notes in Computer Science, vol. 3546, pp. 1100–1109. Springer, New York (2005)
11. Davida, G.I., Frankel, Y., Matt, B.J.: On enabling secure applications through off-line biometric identification. In: Proceedings of the IEEE 1998 Symposium on Security and Privacy, pp. 148–157, Oakland, CA (1998)
12. Monrose, F., Reiter, M.K., Wetzel, S.: Password hardening based on keystroke dynamics. Int. J. Inform. Secur. **1**(2), 69–83 (2002)
13. Hao, F., Anderson, R., Daugman, J.: Combining Crypto with Biometrics Effectively. IEEE Trans. Comput. **55**(9), 1081–1088 (2006)
14. Bringer, J., Chabanne, H., Cohen, G., Kindarji, Z'emor, G.: Optimal iris fuzzy sketches. In: IEEE First International Conference on Biometrics: Theory, Applications, and Systems, BTAS'07, Washington, DC, 27–29 Sept, (2007)
15. Nandakumar, K., Jain, A.K., Pankanti, S.C.: Fingerprint-based Fuzzy Vault: Implementation and Performance. IEEE Trans. Inform. Forensics Secur. **2**(4), 744–757 (2007)
16. Draper, S.C., Khisti, A., Martinian, E., Vetro, A., Yedidia, J.S.: Using Distributed Source Coding to Secure Fingerprint Biometrics. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol. 2, pp. 129–132 (2007)
17. Chang, E.-C., Shen, R., Teo, F.W.: Finding the Original Point Set Hidden among Chaff. In: Proceedings of the 2006 ACM Symposium on Information, computer and communications security. ASIACCS'06, Taipei, Taiwan, pp. 182–188 Sept, (2006)
18. Delvaux, N., Bringer, J., Grave, J., Kratsev, K., Lindeberg, P., Midgren, J., Breebaart, J., Akkermans, T., van der Veen, M., Veldhuis, R., Kindt, E., Simoens, K., Busch, C., Bours, P., Gafurov, D., Yang, B., Stern, J., Rust, C., Cucinelli, B., Skepastianos, D.: Pseudo identities based on fingerprint characteristics. In: IEEE fourth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2008), August 15–17, Harbin, China (2008)

Enhancement

► Biometric Algorithms

Enrollment

Enrollment is the procedure in a biometric system, in which a subject, i.e., an enrollee, presents one or more biometric data samples to the system, and the system then generates from the data biometric templates for future use of biometric matching.

► Biometrics, Overview

Enrollment Time

The time needed to enroll a subject into a biometric system. Enrollment means the process of collecting biometric samples from a subject as well as the subsequent feature extraction to create that person's reference templates representing his identity and finally storing them into the database. Only enrolled subjects can be recognized by the system.

► Performance Evaluation, Overview

Enrollment Transaction Duration

IBG defines the term Enrollment Transaction Duration as the median time duration required for enrollment process of the system. This includes the time for the test subject to align himself or herself with the acquisition device, all biometric sample presentations, the intervals between the first and the second instance enrolment, and the template generation processing time.

► Finger Vein Reader

Ensemble Learning

Zhi-Hua Zhou
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

Synonyms

Committee-based learning; Multiple classifier systems; Classifier combination

Definition

Ensemble learning is a machine learning paradigm where multiple learners are trained to solve the same problem. In contrast to ordinary machine learning approaches which try to learn *one* hypothesis from training data, ensemble methods try to construct a *set* of hypotheses and combine them to use.

Introduction

An ensemble contains a number of learners which are usually called *base learners*. The ▶ generalization ability of an ensemble is usually much stronger than that of base learners. Actually, ensemble learning is appealing because that it is able to boost *weak learners* which are slightly better than random guess to *strong learners* which can make very accurate predictions. So, "base learners" are also referred as "weak learners". It is noteworthy, however, that although most theoretical analyses work on weak learners, base learners used in practice are not necessarily weak since using not-so-weak base learners often results in better performance.

Base learners are usually generated from training data by a *base learning algorithm* which can be decision tree, neural network or other kinds of machine learning algorithms. Most ensemble methods use a single base learning algorithm to produce *homogeneous* base learners, but there are also some methods which use multiple learning algorithms to produce *heterogeneous* learners. In the latter case there is no single base learning algorithm and thus, some people prefer calling the learners *individual learners* or *component learners* to "base learners", while the names "individual learners" and "component learners" can also be used for homogeneous base learners.

It is difficult to trace the starting point of the history of ensemble methods since the basic idea of deploying multiple models has been in use for a long time, yet it is clear that the hot wave of research on ensemble learning since the 1990s owes much to two works. The first is an applied research conducted by Hansen and Salamon [1] at the end of 1980s, where they found that predictions made by the combination of a set of classifiers are often more accurate than predictions made by the best single classifier. The second is a theoretical research conducted in 1989, where Schapire [2] proved that *weak learners* can be boosted to *strong learners*, and the proof resulted in Boosting, one of the most influential ensemble methods.

Constructing Ensembles

Typically, an ensemble is constructed in two steps. First, a number of base learners are produced, which can be generated in a *parallel* style or in a *sequential* style where the generation of a base learner has influence on the generation of subsequent learners. Then, the base learners are combined to use, where among the most popular combination schemes are *majority voting* for classification and *weighted averaging* for regression.

Generally, to get a good ensemble, the base learners should be as more accurate as possible, and as more diverse as possible. This has been formally shown by Krogh and Vedelsby [3], and emphasized by many other people. There are many effective processes for estimating the *accuracy* of learners, such as ▶ cross-validation, hold-out test, etc. However, there is no rigorous definition on what is intuitively perceived as *diversity*. Although a number of diversity measures have been designed, Kuncheva and Whitaker [4] disclosed that the usefulness of existing diversity measures in constructing ensembles is suspectable. In practice, the diversity of the base learners can be introduced from different channels, such as subsampling the training examples, manipulating the attributes, manipulating the outputs, injecting randomness into learning algorithms, or even using multiple mechanisms simultaneously. The employment of different base learner generation processes and/or different combination schemes leads to different ensemble methods.

There are many effective ensemble methods. The following will briefly introduce three representative methods, *Boosting* [2, 5], *Bagging* [6] and *Stacking*

[7]. Here, binary classification is considered for simplicity. That is, let $\mathcal{X}$ and $\mathcal{Y}$ denote the instance space and the set of class labels, respectively, assuming $\mathcal{Y} = \{-1, +1\}$. A training data set $\mathcal{D} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_m, y_m)\}$ is given, where $\mathbf{x}_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$ (i=1,…,m).

Boosting is in fact a family of algorithms since there are many variants. Here, the most famous algorithm, AdaBoost [5], is considered as an example. First, it assigns equal weights to all the training examples. Denote the distribution of the weights at the t-th learning round as D_t. From the training data set and D_t the algorithm generates a base learner $h_t : \mathcal{X} \to \mathcal{Y}$ by calling the base learning algorithm. Then, it uses the training examples to test h_t, and the weights of the incorrectly classified examples will be increased. Thus, an updated weight distribution D_{t+1} is obtained. From the training data set and D_{t+1} AdaBoost generates another base learner by calling the base learning algorithm again. Such a process is repeated for T times, each of which is called a *round*, and the final learner is derived by weighted majority voting of the T base learners, where the weights of the learners are determined during the training process. In practice, the base learning algorithm may be a learning algorithm which can use weighted training examples directly; otherwise the weights can be exploited by sampling the training examples according to the weight distribution D_t. The pseudo-code of AdaBoost is shown in Fig. 1.

Bagging [6] trains a number of base learners each from a different *bootstrap sample* by calling a base learning algorithm. A bootstrap sample is obtained by subsampling the training data set with replacement, where the size of a sample is as the same as that of the training data set. Thus, for a bootstrap sample, some training examples may appear but some may not, where the probability that an example appears at least once is about 0.632. After obtaining the base learners, Bagging combines them by majority voting and the most-voted class is predicted. The pseudo-code of Bagging is shown in Fig. 2. It is worth mentioning that a variant of Bagging, *Random Forests* [8], has been deemed as one of the most powerful ensemble methods up to date.

In a typical implementation of Stacking [7], a number of first-level individual learners are generated from the training data set by employing different learning algorithms. Those individual learners are then combined by a second-level learner which is called as *meta-learner*. The pseudo-code of Stacking is shown in Fig. 3. It is evident that Stacking has close relation with information fusion methods.

Generally speaking, there is no ensemble method which outperforms other ensemble methods consistently. Empirical studies on popular ensemble methods can be found in many papers such as [9, 10, 11]. Previously, it was thought that using more base learners will lead to a better performance, yet Zhou et al. [12] proved the "many could be better than all" theorem which indicates that this may not be the fact. It was shown that after generating a set of base learners, selecting some base learners instead of using all of them to compose an ensemble is a better choice. Such ensembles are called *selective ensembles*.

Input: Data set $\mathcal{D} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_m, y_m)\}$;
Base learning algorithm $\mathcal{L}$;
Number of learning rounds T.

Process:
$D_1(i) = 1/m$. % Initialize the weight distribution
for $t = 1, \cdots, T$:
$h_t = \mathcal{L}(\mathcal{D}, D_t)$; % Train a base learner h_t from $\mathcal{D}$ using distribution D_t
$\epsilon_t = \Pr_{i \sim D_i}[h_t(\boldsymbol{x}_i \neq y_i)]$; % Measure the error of h_t
$\alpha_t = \frac{1}{2} \ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$; % Determine the weight of h_t
$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} \exp(-\alpha_t) & \text{if } h_t(\boldsymbol{x}_i) = y_i \\ \exp(\alpha_t) & \text{if } h_t(\boldsymbol{x}_i) \neq y_i \end{cases}$
$= \frac{D_t(i) \exp(-\alpha_t y_i h_t(\boldsymbol{x}_i))}{Z_t}$ % Update the distribution,where Z_t is a normalization
% factor which enables D_{t+1} to be a distribution
end.

Output: $H(\boldsymbol{x}) = \text{sign}(f(\boldsymbol{x})) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(\boldsymbol{x})\right)$

Ensemble Learning. **Figure 1** The AdaBoost algorithm.

Input: Data set $\mathcal{D} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_m, y_m)\}$;
Base learning algorithm $\mathcal{L}$;
Number of learning rounds T.

Process:
for $t = 1, \cdots, T$:
$\mathcal{D}_t = Bootstrap(\mathcal{D})$; % Generate a bootstrap sample from $\mathcal{D}$
$h_t = \mathcal{L}(\mathcal{D}_t)$ % Train a base learner h_t from the bootstrap sample
end.

Output: $H(\boldsymbol{x}) = \text{argmax}_{y \in \mathcal{Y}} \sum_{t=1}^{T} 1(y = h_t(\boldsymbol{x}))$ % the value of $1(a)$ is 1 if a is $true$ and 0 otherwise

Ensemble Learning. Figure 2 The Bagging algorithm.

Input: Data set $\mathcal{D} = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_m, y_m)\}$;
First-level learning algorithms $\mathcal{L}_1, \cdots, \mathcal{L}_T$;
Second-level learning algorithm $\mathcal{L}$.

Process:
for $t = 1, \cdots, T$:
$h_t = \mathcal{L}_t(\mathcal{D})$ % Train a first-level individual learner h_t by applying the first-level
end; % learning algorithm $\mathcal{L}_t$ to the original data set $\mathcal{D}$
$\mathcal{D}' = \emptyset$; % Generate a new data set
for $i = 1, \cdots, m$:
for $t = 1, \cdots, T$:
$z_{it} = h_t(\boldsymbol{x}_i)$ % Use h_t to classify the training example $\boldsymbol{x}_i$
end;
$\mathcal{D}' = \mathcal{D}' \cup \{((z_{i1}, z_{i2}, \cdots, z_{iT}), y_i)\}$
end;
$h' = \mathcal{L}(\mathcal{D}')$. % Train the second-level learner h' by applying the second-level
% learning algorithm $\mathcal{L}$ to the new data set $\mathcal{D}'$

Output: $H(\boldsymbol{x}) = h'(h_1(\boldsymbol{x}), \cdots, h_T(\boldsymbol{x}))$

Ensemble Learning. Figure 3 The Stacking algorithm.

It is worth mentioning that in addition to classification and regression, ensemble methods have also been designed for clustering [13] and other kinds of machine learning tasks.

Why Ensembles are Superior to Singles

To understand that why the generalization ability of an ensemble is usually much stronger than that of a single learner, Dietterich [14] gave three reasons by viewing the nature of machine learning as searching a hypothesis space for the most accurate hypothesis. The first reason is that, the training data might not provide sufficient information for choosing a single best learner. For example, there may be many learners perform equally well on the training data set. Thus, combining these learners may be a better choice. The second reason is that, the search processes of the learning algorithms might be imperfect. For example, even if there exists a unique best hypothesis, it might be difficult to achieve since running the algorithms result in sub-optimal hypotheses. Thus, ensembles can compensate for such imperfect search processes. The third reason is that, the hypothesis space being searched might not contain the true target function, while ensembles can give some good approximation. For example, it is well-known that the classification boundaries of decision trees are linear segments parallel to coordinate axes. If the target classification boundary is a smooth diagonal line, using a single decision tree cannot lead to a good result yet a good approximation can be achieved by combining a set of decision trees. Note that those are intuitive instead of rigorous theoretical explanations.

There are many theoretical studies on famous ensemble methods such as Boosting and Bagging, yet it is far from a clear understanding of the underlying mechanism of these methods. For example, empirical observations show that Boosting often does *not*

suffer from ▶ overfitting even after a large number of rounds, and sometimes it is even able to reduce the generalization error after the training error has already reached zero. Although many researchers have studied this phenomenon, theoretical explanations are still in arguing.

The ▶ bias-variance decomposition is often used in studying the performance of ensemble methods [9, 12]. It is known that Bagging can significantly reduce the variance, and therefore it is better to be applied to learners suffered from large variance, e.g., unstable learners such as decision trees or neural networks. Boosting can significantly reduce the bias in addition to reducing the variance, and therefore, on weak learners such as decision stumps, Boosting is usually more effective.

Applications

Ensemble learning has already been used in diverse applications such as optical character recognition, text categorization, face recognition, computer-aided medical diagnosis, gene expression analysis, etc. Actually, ensemble learning can be used wherever machine learning techniques can be used.

Summary

Ensemble learning is a powerful machine learning paradigm which has exhibited apparent advantages in many applications. By using multiple learners, the generalization ability of an ensemble can be much better than that of a single learner. A serious deficiency of current ensemble methods is the lack of comprehensibility, i.e., the knowledge learned by ensembles is not understandable to the user. Improving the comprehensibility of ensembles [15] is an important yet largely understudied direction. Another important issue is that currently no diversity measures is satisfying [4] although it is known that diversity plays an important role in ensembles. If those issues can be addressed well, ensemble learning will be able to contribute more to more applications.

Related Entries

▶ AdaBoost
▶ Classifier Design
▶ Multiple Experts
▶ Machine-Learning
▶ Multiple Classifier Systems
▶ Probability Distribution

References

1. Hansen, L.K., Salamon, P.: Neural network ensembles. IEEE Trans. Pattern Analy. Mach. Intell. **12**(10), 993–1001 (1990)
2. Schapire R.E.: The strength of weak learnability. Mach. Learn. **5**(2), 197–227 (1990)
3. Krogh, A., Vedelsby, J.: Neural network ensembles, cross validation, and active learning. In: Tesauro, G. Touretzky, D.S. Leen T.K. (eds.) Advances in Neural Information Processing Systems 7, pp. 231–238. MIT, Cambridge, MA (1995)
4. Kuncheva, L.I., Whitaker, C.J.: Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Mach. Learn. **51**(2), 181–207 (2003)
5. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to Boosting. J. Comput. Syst. Sci. **55**(1), 119–139 (1997)
6. Breiman, L.: Bagging predictors. Mach. Learn. **24**(2), 123–140 (1996)
7. Wolpert, D.H.: Stacked generalization. Neural Networks **5**(2), 241–260 (1992)
8. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
9. Bauer, E., Kohavi, R.: An empirical comparison of voting classification algorithms: Bagging, Boosting, and variants. Mach. Learn. **36**(1–2), 105–139 (1999)
10. Ting, K.M., Witten, I.H.: Issues in stacked generalization. J. Artif. Intell. Res. **10**, 271–289 (1999)
11. Opitz, D., Maclin, R.: Popular ensemble methods: An empirical study. J. Artif. Intell. Res. **11**, 169–198 (1999)
12. Zhou, Z.H., Wu, J., Tang, W.: Ensembling neural networks: Many could be better than all. Artif. Intell. **137**(1–2), 239–263 (2002)
13. Strehl, A., Ghosh, J.: Cluster ensembles - a knowledge reuse framework for combining multiple partitionings. J. Mach. Learn. Res. **3**, 583–617 (2002)
14. Dietterich, T.G.: Machine learning research: Four current directions. AI Mag. **18**(4), 97–136 (1997)
15. Zhou, Z.H., Jiang, Y., Chen, S.F.: Extracting symbolic rules from trained neural network ensembles.. AI Commun. **16**(1), 3–15 (2003)

Entropy, Biometric

Biometric entropy describes the inherent variability in biometric samples in the population. It can also be understood as the information content of biometric samples is related to many questions in biometric

technology. For example, one of the most common biometric questions is that of uniqueness (e.g., "are fingerprints unique?"). Such a measure is important for the performance of biometric system, as a measure of the strength of biometric cryptosystems and for privacy measures. It also is relevant for applications such as biometric fusion, where one would like to quantify the biometric information in each system individually, and the potential gain from fusing the systems. Many approaches have been taken to measure biometric entropy, like Wayman (2004) introduced a statistical approach to measure the separability of Gaussian feature distributions using a "cotton ball model". Daugman (2003) developed "discrimination entropy" to measure the information content of iris images. This value has the advantage that it is calculated directly from the match score distributions, but how it relates to traditional measures of entropy is not clear. Golfarelli et al. (1997) showed that the most commonly used feature representations of hand geometry and face biometrics have a limited number of distinguishable patterns, as measured by a theoretical estimate of the equal error rate. Penev et al. (2000) determined the dimensionality of the PCA subspace necessary to characterize the identity information in faces. Adler et al. (2005) defined biometric entropy as the "decrease in uncertainty about the identity of a person due to a set of biometric measurements," and expressed in terms of the relative entropy D(pkq) between the population (inter-class) feature distribution q and the individual (intra-class) distribution p. Biometric entropy still does not have a well accepted definition. Additionally, all proposed schemes measure the information content of a feature representation, and not that of the biometric sample itself.

► Security and Liveness, Overview

ePassport

ePassport = e-Passport or IC Passport. Biometrically enabled passports that meet the requirements of a facial biometric, which can be captured from a submitted photograph.

► Photography for Face Image Data

Epidermis

Epidermis is the uppermost layer of the skin. Its thickness varies depending upon the location of the skin. Generally it is found to be 0.5 nm on the eyelids (thinnest) and 1.5 nm at palm and sole (thickest). It consists of five layers named:

1. Stratum germinatum
2. Stratum granulosum
3. Stratum spinosum
4. Stratum licidum
5. Stratum corneum

► Anatomy of Hand
► Skin Spectroscopy

Ergonomic Design for Biometric Systems

Eric P. Kukula, Stephen J. Elliott
Purdue University, West Lafayette, IN, USA

Synonyms

Human-Biometric Sensor Interaction (HBSI); Human-Computer Interaction (HCI); Human Factors; Usability

Definition

Biometric ergonomic design is the area of research that examines how humans interact and use biometric sensors, devices, interfaces, and systems. The purpose is to understand the physical and cognitive human-biometric sensor interaction to improve the system design and overall performance of a biometric system.

Introduction

Biometric ► ergonomic design is an emerging interdisciplinary research area in biometrics that focuses on the ► interaction between the user and the biometric

system to better understand issues and errors, users knowingly or unknowingly generate when attempting to use a biometric system. This research area attempts to understand what tasks, movements, and behaviors users execute when encountering different biometric modalities. This area presents a challenge for the biometrics community – while the algorithms are continually improving, there are still individuals who cannot successfully interact with the biometric sensor(s). It is essential that designers continue examining biometric devices, process, or systems to ensure they accomodate the focal point of any biometric systems, *the human*. Adapting devices, processess or systems to the *human* can increase usability by minimizing errors during presentation and acquistion of the biometric characteristics to the sensor through better design, instruction, or system feedback.

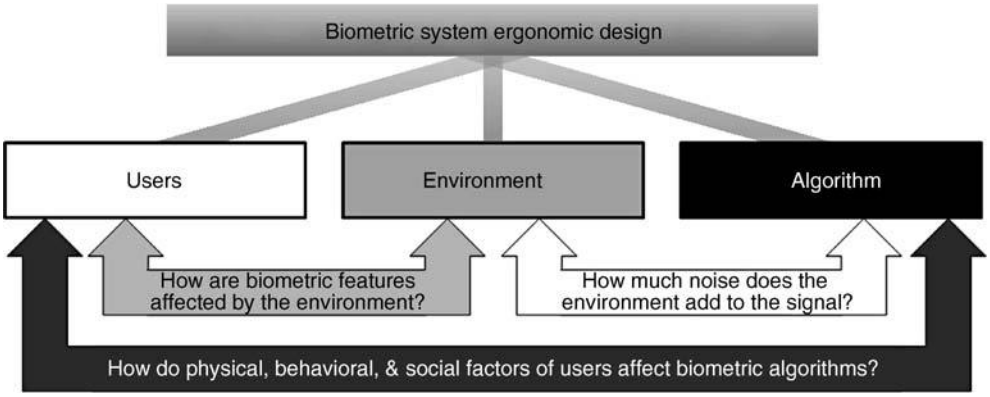
Traditional approaches to evaluate the performance of a biometric system have been system-level, meaning that evaluators and designers are more interested in system reported error rates, some of which include: the *Failure to Enroll (FTE) rate*, *Failure to Acquire (FTA) rate*, *False Accept Rate (FAR)*, and *False Reject Rate (FRR)*. Traditional performance evaluations have worked well to evaluate emerging technologies, new biometric modalities, and algorithm revisions, which are typically associated with *technology performance evaluations*. Moreover, since biometrics entered the commercial marketplace, most research has been dedicated to the development in three areas: (1) improving performance, (2) increasing throughput, and (3) decreasing the size of the sensor or hardware device. Limited research has focused on ergonomic design and usability issues, which relate to how users interact and use biometric devices. No standard activities have focused on ergonomic design or usability issues with biometrics, although standard testing and evaluation protocols do exist, specifically – ISO 19795-1: Technology Testing [1], ISO 19795-2: Scenario Testing [2], and ISO TR19795-3: Modality-Specific Testing [3].

While early research has been concerned with the design, development, and testing of biometric systems and algorithms, recent research has attributed human physical, behavioral, and social factors to affect the performance of the overall biometric system. Moreover, these factors are of utmost importance when conducting *scenario* and *operational performance evaluations*, as they are the last line of defense between the laboratory and the commercial marketplace to understand how a biometric system performs in a particular environment or with a specific set of users. Therefore as the community continues to learn more about the different biometric modalities and systems, as well as how users interact with them, performance from both the system and user perspectives must be fully understood to make further improvements to the biometric sensor, algorithm, and design of future user interfaces.

Biometric Properties and Ergonomic Implications

Biometric modalities are classified as physiological, behavioral, or a combination of the two. In addition, they are classified according to five desirable properties, outlined by Clarke [4], and amended by numerous others. Desirable properties of biometric characteristics are that they offer: (1) universality – available in all people, (2) invariant – features extracted are non-changing, (3) high intra-class variability – features extracted are distinct for each user; (4) acceptability – characteristic of suitability for use by everyone, and (5) extractability – a sensor can extract the features presented in a repeatable manner. Although commonly described in the literature as the ideal characteristics of the biometric, each must overcome challenges. Herein lies one of the challenges associated with large-scale deployment of biometrics and the purpose behind research in this area – the majority of biometrics are challenged to satisfy all these five categories.

To better understand the importance of ergonomics in biometrics, the authors pose the question: *what affects biometric system performance*? Generalizing the issues that can be linked to many performance failures into three divisions, bins for users (physical, behavioral, and social factors), the environment, and matching algorithms emerge. While it is important to understand each group when designing a biometric system, the inter-relationship between the groups also impacts biometric performance, which is illustrated in Fig. 1. First, the user-environment relationship impacts performance. For example, climatic or work conditions may require individuals to wear personal protective equipment (PPE), which not only limits biometric modalities that can be deployed, but may also occlude the biometric characteristics from being successfully



Ergonomic Design for Biometric Systems. **Figure 1** Issues that affect biometric system performance and the relationship with ergonomics.

acquired in the first place, such as the case in safety glasses for iris recognition. In addition, atmospheric conditions such as temperature and humidity can impact the skin, affecting the acquisition for some modalities. Second is the environment and interrelationship of algorithm. Examples of this include ambient noise for voice recognition and illumination or busy backgrounds for face recognition. Third is the relationship between users and algorithms. First, physiological factors such as skin moisture, elasticity, age, and color can affect performance of algorithms. Secondly, behavioral factors such as finger preference can impact performance. For example, individuals of Asian descent prefer to use the little finger for fingerprint recognition, but it is documented in the literature [5, 6] that the little finger is the worst performing finger. Lastly, social preferences or factors such as hair length or the wearing of head coverings can impact face and iris recognition due to the occlusion of necessary features. While the literature has investigated some of the aforementioned items, more research is needed in these areas. However, there is also an interaction between the three clusters as indicated in the research conducted by the Kukula, et al. [7, 8], but it has not been thoroughly investigated.

It is well documented in the literature that image quality affects the biometric matching algorithm. Yao et al. [9] stated that "in a deployed system, the poor acquisition of samples perhaps constitutes the single most important reason for high false reject/accept rates" and further discussed that there are two solutions for reducing poor images. First, one can model and weight all adverse situations for the feature extraction and matching system. Second, "one can try to dynamically and interactively obtain a desirable input sample." Improving the ergonomic design of biometric systems is one method to dynamically "modify" the input sample through improved usability of biometric devices, processes, and systems.

Common Design Concerns

Biometric systems are heavily dependent on the sensor to acquire the sample, segment it, and extract features from samples for the matcher to determine the correct response. By observing how users interact with biometric sensors, several design issues are apparent but could be resolved by integrating knowledge of industrial design, ergonomics and ▶ human factors, and ▶ usability. Rubin [10] discusses five reasons why products and systems are difficult to use. The main problem is that the emphasis and focus has been on the machine/system and not on the end user during development. Common design misconceptions are:

- Humans are flexible and will adjust to a product or device
- Engineers work well with technology but not with people
- Engineers are hired to solve technology problems and not people skills
- Designers create products for users like themselves in terms of both usage and level of knowledge [10]

The following factors are true within the context of biometric system design. Humans will adapt to the sensor and/or system. Many times, biometric systems or sensors are not tested on sufficiently large

numbers of the general populations, namely due to the cost of doing so. Moreover, the biometric community may test the algorithms exhaustively off-line, using pre-collected images, but lapse on collecting images with a new sensor to examine how the user interacts with the system or device.

According to Smith [11], some members of the Human-Computer Interaction (HCI) community believe that interfaces of security systems do not reflect good thinking in terms of creating a system that is easy to use, while maintaining an acceptable level of security (p. 75). Moreover, according to Adams and Sasse [12], security systems are one of the last areas to embrace user-centered design and training as essential. This is also true for biometrics as Coventry et al. [13] stated the Human-Computer Interaction (HCI) community has had limited involvement in the design or evaluation of biometric systems.

Human-Biometric Sensor Interaction (HBSI)

The authors have been researching this area for over four years. Results of this research have produced a new conceptual model, which is shown in Fig. 3. This model combines literature and models from biometrics, ergonomics, and usability (Fig. 2). The conceptual model that examines biometric system ergonomic design is called the Human-Biometric Sensor Interaction, or HBSI. The three fields of biometrics, ergonomics, and usability are arranged within the model to show the relationship of the human, biometric sensor, and the biometric system. Each of the relationships poses a different set of design or research questions, which will now be discussed.

Human-Biometric Sensor

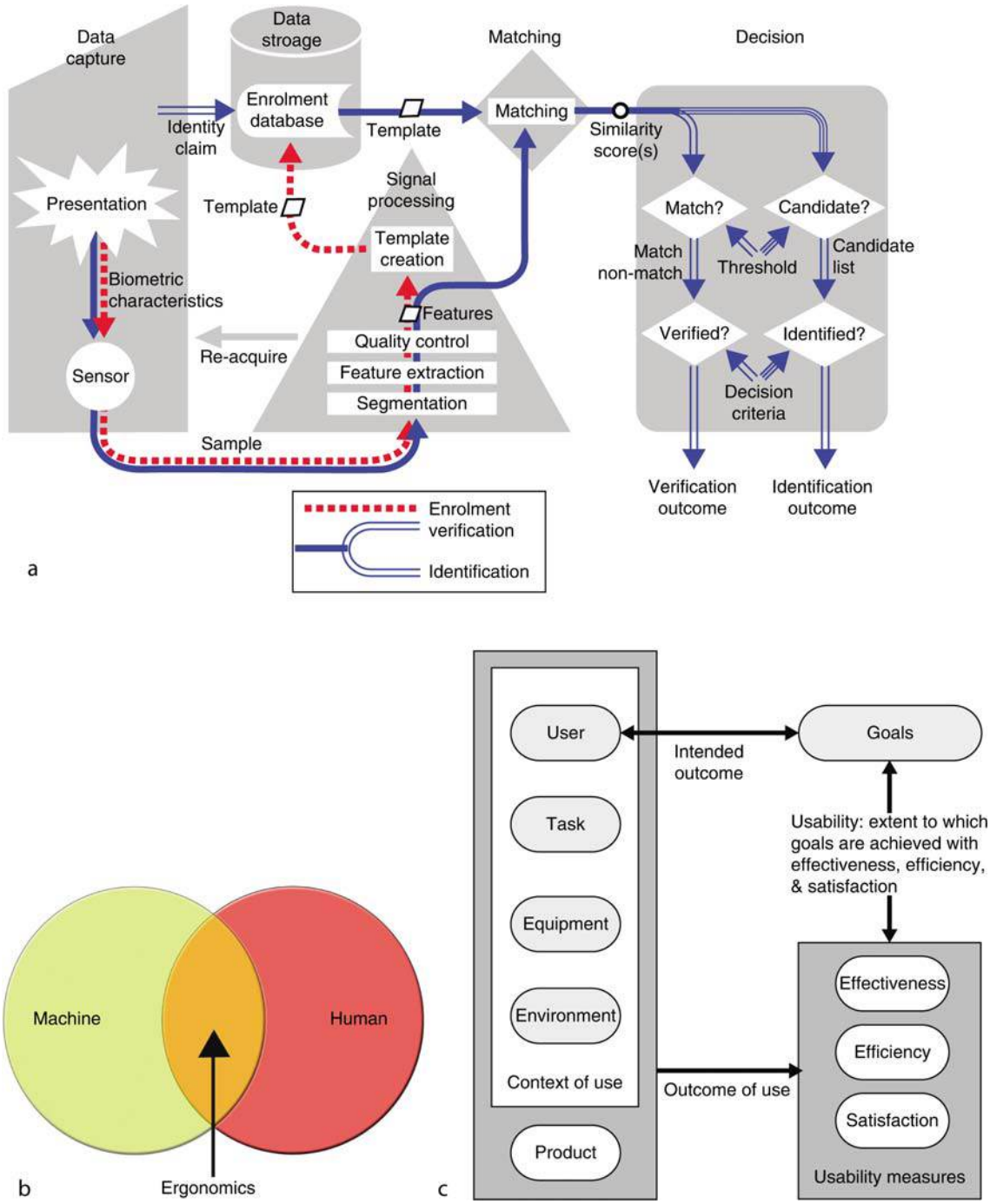
The human and sensor components of the HBSI model are similar to Tayyari and Smith's [14] human-machine interaction model. Much like the traditional model, the human and biometric sensor components look to achieve the optimal relationship between humans and a biometric sensor in a particular environment. The Human-Biometric sensor relationship parallels the presentation silo of the general biometric model, and is often overlooked during the design of the biometric system. Applying an ergonomic approach during the design of the biometric sensor we can fit the sensors to the majority of users, as opposed to forcing users to interact with difficult and uncomfortable biometric sensors. Applying ergonomic approaches such as ▶ user-centered design, biometric sensors, interfaces, and systems can be designed based on the user's physical and mental states to allow the users to complete the task that the biometric system is asking for, most efficiently.

Human-Biometric System

The human and biometric system components of the HBSI model are arranged to accommodate the way that biometric sensors, software, and implementations are presented to users. Not only a biometric sensor must be designed so that a user can interact with it in a repeatable fashion, but also the sensor(s), software, and the way the entire "system" is packaged must be usable. Usability according to ISO 9241-11 [15] is segmented into three factors: effectiveness, efficiency, and satisfaction. Each of the three metrics is distinctively different and important to understand. System designers must take into consideration the goals of the system. Every biometric system will be designed for a different purpose, thus a balance must be attained between effectiveness, efficiency, and satisfaction. First, biometric systems must be effective, meaning users are able to interact, use, and complete the desired tasks without too much effort, which can also cause throughput issues if people get "lost" in the system and require administrator intervention, which also comes with a cost. Second, biometric systems must be efficient, meaning users must be able to accomplish the tasks easily and in a timely manner. Again, if users require intervention, the cost of staffing becomes burdensome. Third, users must like, or be satisfied, with the biometric system, or will discontinue use and find alternative methods to accomplish the task.

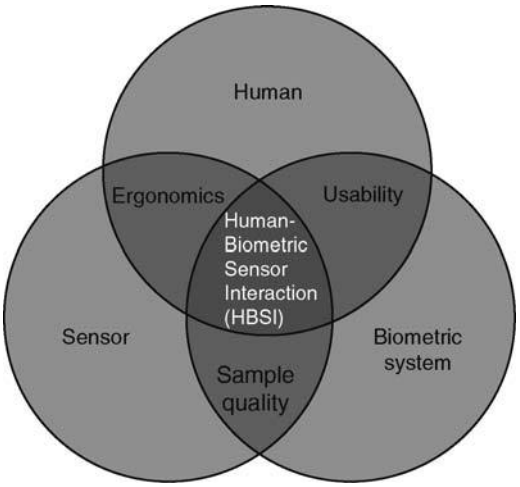
Sensor-Biometric System

As mentioned in the previous two sections, users must be able to interact with a biometric sensor or device in a consistent manner over time; however, users must also find the entire biometric system usable. To enable this, the third relationship of the HBSI conceptual



Ergonomic Design for Biometric Systems. Figure 2 General biometric model **(a)** [1], general ergonomic model **(b)** [14], and general usability model **(c)** [15].

model emerges, i.e., the sensor-biometric system measured by image quality. Image quality is the important link between these two components because the image or sample acquired by the biometric sensor must contain the characteristics or features needed by the biometric system to enroll or match a user in the biometric system. So not only does the human-sensor relationship needs to be functional and the human-biometric system needs to be usable, but also the sensor-biometric system needs to be efficient. This occurs only if the sensor captures and passes usable features onto the biometric system.



Ergonomic Design for Biometric Systems. **Figure 3** The Human-Biometric Sensor Interaction or HBSI model.

The Human-Biometric Sensor Interaction

The combination of components and relationships in the model form the Human-Biometric Sensor Interaction. Each component that is in the HBSI model has been shown to impact results in previous experiments from the respective field from which it was adapted from. Since the conceptual model is derived from different fields, each component usability, ergonomics, and biometrics produces a unique output. Thus, the final determination of the results is dependent upon the goals, objectives, and criteria the researcher, designer, or engineer is seeking, which is in-line with the ergonomics, usability, and design literature.. As work in the area of biometric system ergonomic design is limited, the HBSI model provides the biometrics community more insight and considerations needed for designing biometric systems and their corresponding devices, as well as metrics to evaluate the components outside traditional biometric testing and evaluation.

Literature

Seminal research and publication in the area of usability and accessibility, which was concerned with biometric system ergonomic design, were pioneered by the User Research Group at National Cash Register (NCR). Some of their research findings that would impact biometric system design can be seen in the results of one experiment, which revealed that successful verification was not affected by the type of instruction and feedback received. Furthermore, the results also revealed some users have problems that cannot be solved through instruction, training, or feedback. A possible explanation could be the biometric system ergonomic design and placement of the sensor and the human-biometric sensor interaction. Please refer to a book chapter written by Coventry [16] for more information and relevant citations of work conducted by the User Research Group at NCR.

Two other groups that have been actively researching and publishing in this area are the NIST Biometrics and Usability Group [17] and Purdue University's Biometric Standards, Performance, & Assurance Laboratory [18]. Please refer to the respective references for the latest research, publications, and presentations in the area of biometric system ergonomic design. At the time of writing, research in this area has investigated ten print fingerprint capture scanner height and angle, hand geometry device height, ▶ habituation, applied finger force on a fingerprint sensor, and usability of small-area and swipe-based fingerprint sensors, image quality evaluations, instruction and feedback mechanisms, as well as health and safety perceptions of biometric devices. Lastly, the United Kingdom Home Office Identity and Passport Service has also published reports based on their biometric trials and implementations which discuss biometric usability and ergonomic design [19]. Maple and Norrington [20] reported one particular trial of the United Kingdom's Passport Service Trial Program and its usability and found issues with each of the three evaluated biometric systems: fingerprint, face, and iris recognition systems.

Summary

This entry discussed the effect human interaction has on biometric system performance to outline the impact biometric system ergonomic design can have on the overall performance of a biometric system. The entry has outlined the origins of the Human-Biometric Sensor Interaction model; including relevant work and models in the fields of ergonomics, user-centered design, usability, and HCI. In addition, this entry has discussed how the fields that form the HBSI model not only relate to biometrics, but can be integrated into the design of biometric devices and systems to create more usable devices and systems, with the goal of lowering acquisition,

enrollment, and matching failures. However, further understanding in the area of biometric system ergonomic design and its impact on biometrics is needed to meet this goal.

The authors are not alone in their thoughts and opinions that continued research is needed in the area of biometric system ergonomic design. As Smith [11] stated that some members of the HCI community believe that interfaces of security systems do not reflect good thinking in terms of creating a system that is easy to use, while maintaining an acceptable level of security. Moreover Adams and Sasse discussed the fact that security systems are one of the last areas to embrace user-centered design and training as essential [12]. Lastly, Maple and Norrington [20], noted three observations that align with the objective for continued investigation in biometric system ergonomic design:

- People have different cognitive abilities,
- People have different physical characteristics and interact differently with equipment, and
- People have different sensory abilities and will perceive biometric sensors and systems differently.

As the biometrics community continues to develop biometric systems and deployments become more pervasive, the evaluation of the biometric system and the respective human-biometric sensor interaction will continue to gain traction.

Related Entries

- ▶ Accessibility
- ▶ Attempt
- ▶ Failure to Acquire (FTA)
- ▶ Failure to Enroll (FTE)
- ▶ Usuability

References

1. International standards organization, information technology – Biometric performance testing and reporting – Part 1: Principles and framework. ISO/IEC: Geneva. p. 56. (2006)
2. International standards organization, information technology – Biometric performance testing and reporting - Part 2: Testing methodologies for technology and scenario evaluation. ISO/IEC: Geneva. p. 48. (2007)
3. International standards organization, Text of DTR 19795-3, Biometric performance testing and reporting – Part 3: Modality-specific testing. ISO/IEC: Geneva. p. 28. (2007)
4. Clarke, R.: Human identification in information systems: Management challenges and public policy issues. Inf. Technol. People **7**(4), 6–37 (1994)
5. Young, M., Elliott, S.: Image Quality and Performance Based on Henry Classification and Finger Location. In IEEE Workshop on Automatic Identification Advanced Technologies. Alghero, Italy (2007)
6. Wayman, J.: Multi-finger penetration rate and ROC variability for automatic fingerprint identification systems. In: Wayman, J. (ed.) National Biometric Test Center Collected Works 1997–2000, pp. 179–190. San Jose, CA (2000)
7. Kukula, E.: Design and Evaluation of the Human-Biometric Sensor Interaction Method. In: Industrial Technology, Vol. 1, Ph.D, pp. 510. Purdue University: West Lafayette (2008)
8. Kukula, E., Elliott, S., Duffy, V.: The effects of human interaction on biometric system performance in 12th International Conference on Human-Computer Interaction and 1st International Conference on Digital-Human Modeling. pp. 903–913, Springer, Beijing, China (2007)
9. Yao, M., Pankanti, S., Haas, N.: Fingerprint quality assessment. In: Ratha, N., Bolle, R. (eds.) Automatic Fingerprint Recognition Systems, pp. 55–66. Springer, New York (2004)
10. Rubin, R.: Handbook of usability testing: How to plan, design, and conduct effective tests. Wiley, New York (1994)
11. Smith, S.: Humans in the loop: Human-computer interaction and security. IEEE Secur. Priv. **1**(3), 75–79 (2003)
12. Adams, A., Sasse, M.: Users are not the enemy: Why users compromise security mechanisms and how to take remedial measures. Commun. ACM **42**(12), 41–46 (1999)
13. Coventry, L., De Angeli, A., Johnson, G.: Usability and biometric verification at the ATM interface. In: Conference on Human Factors in Computing Systems. Ft. Lauderdale, ACM Press, Florida (2003)
14. Tayyari, F., Smith, J.: Occupational ergonomics: Principles and applications. In: Parsaei, H. (ed.) Manufacturing Systems Engineering Series, p. 452. Kluwer Academic Publishers, Norwell (2003)
15. International organization for standardization, ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability. p. 28. (1998)
16. Coventry, L.: Usable biometrics. In: Cranor, L.F., Garfinkel, S. (eds.) Security and usability: Designing secure systems that people can use, pp. 175–198. O'Reilly Media, Inc: Sebastopol, CA (2005)
17. NIST. Biometrics and usability group. Available from: http://zing.ncsl.nist.gov/biousa/index.html (2007). Accessed 30 November 2007
18. Purdue University Biometric Standards Performance & Assurance Laboratory. Human Biometric Sensor Interaction. 2007 [cited 2007 November 30]; Available from: http://www.bspalabs.org/archives/category/research/hbsi
19. Home office identity & passport service. Publications. Available from: http://www.ips.gov.uk/passport/publications-general.asp (2007). Accessed 30 November 2007
20. Maple, C., Norrington, P.: The usability and practicality of biometric authentication in the workplace. In: First International Conference on Availability, Reliability and Security (ARES'06). IEEE, Vienna, Austria (2006)

Ergonomics

Ergonomics is a derivative of the Greek words "ergon," or work, and "nomos," meaning laws. While the term work has been traditionally associated with occupation, a broader sense of the term can be applied to any unplanned activity requiring skill or effort. In 2000, the International Ergonomics Association (IEA) defined ergonomics or human factors as: "The scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance." In design, ergonomics attempts to achieve an optimal relationship between humans and machines in a particular environment. The goal of ergonomics, according to Tayyari and Smith, is to "fit (adapt) work to individuals, as opposed to fitting individuals to the work."

► Ergonomic Design for Biometric System

Error Probability Non-Accumulation

► Score Normalization Rules in Iris Recognition

Evaluation of Gait Recognition

Sudeep Sarkar[1], Zongyi Liu[2]
[1]Computer Science and Engineering, University of South Florida, Tampa, FL, USA
[2]Amazon.com, Seattle, WA, USA

Synonyms

Gait recognition

Definition

Gait recognition refers to automated vision methods that use video of human gait to recognize or to identify a person. Evaluation of gait recognition refers to the benchmarking of progress in the design of gait recognition algorithms on standard, common, datasets.

Introduction

Design of biometric algorithms and evaluation of performance goes hand in hand. It is important to constantly evaluate and analyze progress being at various levels of biometrics design. This evaluation can be of three types: at algorithm-level, at scenario-level, and at operational-level, roughly corresponding to the maturity of the biometric. Given the young nature of gait as a biometric source, relative to the mature biometrics such as fingerprints, current evaluations are necessarily at algorithm-level. The motivation behind algorithm-level evaluations is to explore possibilities, to understand limitations, and to push algorithmic research towards hard problems. Some of the relevant questions are

1. Is progress being made in gait recognition of humans?
2. To what extent does gait offer potential as an identifying biometric?
3. What factors affect gait recognition and to what extent?
4. What are the critical vision components affecting gait recognition from video?
5. What are the strengths and weaknesses of different gait recognition algorithms?

An overview of the current evaluation of gait as a potential biometric is discussed here, with particular emphasis on the progress with respect to the HumanID gait challenge problem that has become the de-facto benchmark. A synthesis of gait recognition performances reported on this dataset and other major ones is provided here, along with some suggestions for future evaluations.

A Panoramic View of Performance

To take stock of the progress made in gait recognition, consider a summary of the identification rates reported in the recent literature on different kinds of publicly available experimental protocols and datasets (>25 persons) such as the CMU-Mobo dataset [1] (indoor,

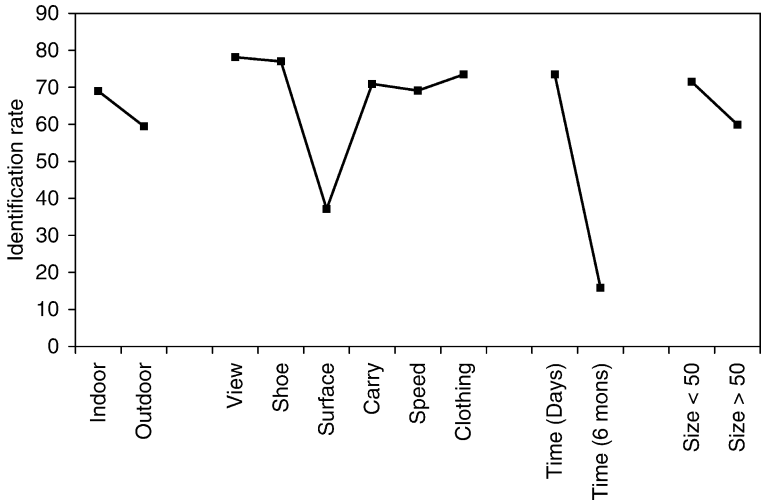
25 subjects), the UMD dataset [2] (outdoor, 55 subjects), the Southampton Large dataset [3] (indoor and outdoor, 115 subjects), the CASIA Gait Dataset [4] (indoor, 124 subjects), and the HumanID Gait Challenge dataset [5] (outdoor, 122 subjects). Figure 1 lists the average identification rates for matching across different conditions, i.e., the ▶ probe and the gallery differed with respect to the indicated ▶ covariate. Of course, the caveat is that the conclusions are conditioned on the kinds of variations of each covariate observed in the respective datasets. Hence, a definitive conclusion is hard to make. However, this kind of summary has some conclusive weight since, it encompasses the findings of multiple research groups. It should provide some directions for focusing future research. The data shows that outdoor gait recognition, recognition across walking surface-type change, and recognition across months are all hard problems. Clothing, footwear, carrying condition, and walking speed does not seem to be hard covariates to overcome. As expected, performance also drops with dataset size, which suggests that it is imperative to demonstrate the efficacy of an idea on as large a dataset as possible.

A deeper look at the performances reported on commonly available datasets, in particular the HumanID gait challenge problem, will form the basis for more definitive conclusions about the progress that is being made.

The HumanID Gait Challenge Problem

The development of gait biometrics is following a path that is somewhat different from other biometrics, for which serious evaluation benchmarks appeared after years of algorithmic development. It was more than 20 years for face recognition, whereas evaluation framework for gait recognition appeared in less than 10 years after the first publication of vision algorithms for gait recognition. Bulk of the research in gait recognition was spurred by the DARPA HumanID at a distance program. The HumanID gait challenge problem was formulated in this program to facilitate objective, quantitative measurement of gait research progress on a large dataset [5]. As of end of 2007, this dataset has been distributed to more than 40 research groups. Many gait recognition research papers report performance on this dataset.

This challenge problem does not just consist of a dataset, but also provides a well-defined experimental



Evaluation of Gait Recognition. **Figure 1** Summary (average) of gait identification rates as reported in the literature for different conditions. The average of the reported rates are listed for different conditions. The first two performance points are average of reported rates on datasets collected indoors and outdoors, respectively. The next six performance points are for matching gait templates across different conditions. For example, "carrying condition" refers to matching gait sequences where the hands of the subjects were free to sequences where the subjects were carrying a briefcase. The "time" condition refers to matching gait templates collected at different times with the time-gap as noted. The size condition refers to number of subjects used in the experiments.

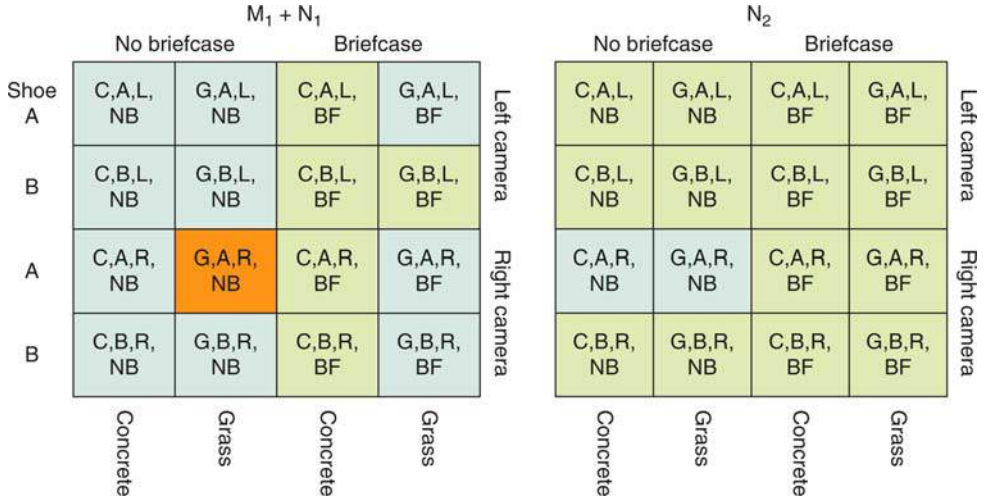
framework for others to follow, along with an established benchmark.

The Dataset

The data was collected outdoors. For each person in the data set, there are combination of as many as five conditions or covariates. The conditions are: (1) two camera angles (L and R), (2) two shoe types (A and B), (3) two surfaces (grass and concrete), (4) with and without carrying a briefcase (B or NB), and (5) two different dates 6 months apart, May and November. The covariates were chosen based on consultation with gait recognition researchers in the Human ID program. These are, of course, not the only variables that can impact gait, but were logistically feasible and likely to impact gait the most. Attempt was made to acquire a person's gait in all possible combinations, and there are up to 32 sequences for some persons. Hence, the full dataset can be partitioned into 32 subsets, one for each combination of the five covariates. The partitioning of the data is visualized in Fig. 2. Each cell refers to an unique combination of view, shoe type, and surface covariates. The smaller arrangement of cells represent the data from repeat subjects. Comparisons between these subsets are used to set up challenge experiments; more on this later. The full data set consists of 1,870 sequences from 122 individuals. This dataset is unique in the number of covariates exercised. It is the only data set to include walking on a grass surface. Figure 3 shows some sample frames from this dataset.

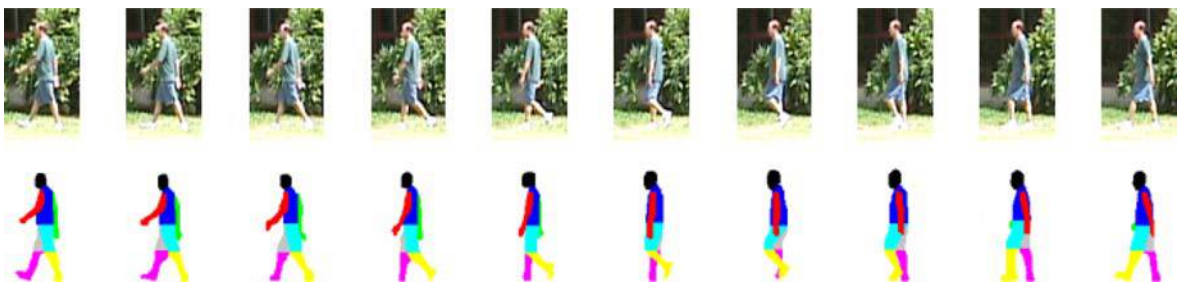
In addition to the raw data sequence, there is an ancillary information associated with the data. First, for each sequence, there is meta-data information about the subject's age, sex, reported height, self reported weight, foot dominance, and shoe information.

Second, for a subset of this dataset, manually created ▶ silhouettes (see Fig. 4) are available. These manual silhouettes should not be used to test any recognition algorithm, but they could be used to build models or to study segmentation errors. More details about the process of creating these manual silhouettes and the quality checks performed can be found in [6]; here are some salient aspects. Seventy one subjects from one of the two collection periods (May collection) were chosen for manual silhouette specification. The sequences corresponding to these subjects were chosen from the (1) gallery set (sequences taken on grass, with shoe type A, right camera view), (2) probe B (on grass, with shoe type B, right camera view), (3) probe D (on concrete, with shoe type A, right camera view), (4) probe H (on grass, with shoe A, right camera view, carrying briefcase), and probe K (on grass, elapsed time). The silhouette in each frame over one walking cycle, of approximately 30–40 image frames was manually



Evaluation of Gait Recognition. Figure 2 Partitioning of the HumanID gait challenge dataset in terms of its covariates, which are coded as follows: C – concrete surface, G – grass surface, A – first shoe type, B – second shoe type, BF – carrying a briefcase, NB – no briefcase, M – data collected in May, N_1 – new subjects in November data, and N_2 – repeat subjects in November. The *shaded cells* are used to design the challenge experiments.

Evaluation of Gait Recognition. Figure 3 Frames from (**a**) the left camera for concrete surface, (**b**) the right camera for concrete surface, (**c**) the left camera for grass surface, (**d**) the right camera for grass surface.



Evaluation of Gait Recognition. Figure 4 *Top row* shows the color images, cropped around the person, for one sequence. The *bottom row* shows the corresponding part-level, manually specified silhouettes.

specified. This cycle was chosen to begin at the right heel strike phase of the walking cycle through to the next right heel strike. Whenever possible, this gait cycle was selected from the same 3D location in each sequence. In addition to marking a pixel as being from the background or subject, more detailed specifications in terms of body parts were marked. The head, torso, left arm, right arm, left upper leg, left lower leg, right upper leg, and right lower leg were explicitly labeled using different colors.

The Challenge Experiments

Along with the dataset, the gait challenge problem includes a definition of 12 challenge experiments (A–L), spanning different levels of difficulty. This provides a common benchmark to compare performance with other algorithms. The experiments are designed to investigate the effect on performance of five factors, i.e., change in viewing angle, change in shoe type, change in walking surfaces (concrete and grass),

carrying or not carrying a briefcase, and temporal differences. The gallery set is common for all the experiments and corresponds to the dark colored cell in Fig. 2. The gallery consists of sequences with the following covariates: Grass, Shoe Type A, Right Camera, No Briefcase, and collected in May along with those from the *new* subjects from November. This set was selected as the gallery because it was one of the largest for a given set of covariates. The experiments differ in terms of the probe sets, which are denoted by the lightly shaded cells. The structure of the 12 probe sets is listed in Table 1. The signatures are the video sequences of gait. The last two experiments study the impact of elapsed time. The elapsed time covariate implicitly includes a change of shoe and clothing because the subjects were not required to wear the same clothes or shoes in both data collections. Because of the implicit change of shoe, it can be safely assumed that a different set of shoes were used in the May and November data collections. This is noted in Table 1 by A/B for shoe type in experiments K and L. The key experiments are those that involve controlled change in just one covariate and are marked with an asterisk in the table. The results from the 12 experiments provide an ordering of difficulty of the experiments.

Baseline Gait Algorithm

The third aspect of the gait challenge problem is a simple but effective ▶ baseline algorithm to provide performance benchmarks for the experiments. Ideally, this should be a combination of "standard" vision modules that accomplishes the task. Drawing from the success of template based recognition strategies in computer vision, a four-part algorithm that relies on silhouette template matching was designed. The first part semi-automatically defines bounding boxes around the moving person in each frame of a sequence. The second part extracts silhouettes from the bounding boxes using expectation maximization based on Mahalanobis distance between foreground and background color model at each pixel. Each silhouette is scaled to a height of 128 pixels and centered (automatically) in each frame along the horizontal direction so that the centerline of the torso is at the middle of the frame.

Evaluation of Gait Recognition. **Table 1** The gallery and probe set specifications for each of gait challenge experiments

Exp.	Probe (surface, shoe, view, carry, elapsed time) (C/G, A/B, L/R, NB/BF, time)	Number of subjects	Difference
A[a]	(G, A, L, NB, M + N_1)	122	V[b]
B[a]	(G, B, R, NB, M + N_1)	54	S[c]
C	(G, B, L, NB, M + N_1)	54	S+V
D[a]	(C, A, R, NB, M + N_1)	121	F[d]
E	(C, B, R, NB, M + N_1)	60	F+S
F	(C, A, L, NB, M + N_1)	121	F+V
G	(C, B, L, NB, M + N_1)	60	F+S+V
H[a]	(G, A, R, BF, M + N_1)	120	B[e]
I	(G, B, R, BF, M + N_1)	60	S+B
J	(G, A, L, BF, M + N_1)	120	V+B
K[a]	(G, A/B, R, NB, N_2)	33	T[f]+S+C[g]
L	(C, A/B, R, NB, N_2)	33	F+T+S+C

The gallery for all of the experiments is (G, A, R, NB, M +N_1) and consists of 122 individuals
[a]Keyexperiments
[b]View
[c]Shoe
[d]Surface
[e]Carry
[f]Elapsed time
[g]Clothing

The third part computes the gait period from the silhouettes. The gait period is used to partition the sequences for spatial-temporal correlation. The fourth part performs spatial-temporal correlation to compute the similarity between two gait sequences.

Let $\mathbf{S_P} = \{\mathbf{S_P}(1), \cdots, \mathbf{S_P}(M)\}$ and $\mathbf{S_G} = \{\mathbf{S_G}(1), \cdots, \mathbf{S_G}(N)\}$, be the probe and the gallery silhouette sequences, respectively. First, the probe (input) sequence is partitioned into subsequences, each roughly over one gait period, N_{Gait}. Gait periodicity is estimated based on periodic variation of the count the number of foreground pixels in the lower part of the silhouette in each frame over time. This number will reach a maximum when the two legs are farthest apart (full stride stance) and drop to a minimum when the legs overlap (heels together stance).

Second, each of these probe subsequences, $\mathbf{S_{Pk}} = \{\mathbf{S_P}(k), \cdots, \mathbf{S_P}(k + N_{\text{Gait}})\}$, is cross correlated with the given gallery sequence, $\mathbf{S_G}$.

$$\text{Corr}(\mathbf{S_{Pk}}, \mathbf{S_G})(l) = \sum_{j=1}^{N_{\text{Gait}}} S(\mathbf{S_P}(k+j), \mathbf{S_G}(l+j)), \quad (1)$$

where, the similarity between two image frames, $S(\mathbf{S_P}(i), \mathbf{S_G}(j))$, is defined to be the Tanimoto similarity between the silhouettes, i.e., the ratio of the number of common pixels to the number of pixels in their union. The overall similarity measure is chosen to be the median value of the maximum correlation of the gallery sequence with each of these probe subsequences. The strategy for breaking up the probe sequence into subsequences allows the algorithm to overcome segmentation errors in some contiguous sets of frames due to some background subtraction artifact or due to localized motion in the background.

$$\text{Sim}(\mathbf{S_P}, \mathbf{S_G}) = \text{Median}_k \left(\max_l \text{Corr}(\mathbf{S_{Pk}}, \mathbf{S_G})(l) \right). \quad (2)$$

The baseline algorithm is parameter free. The algorithm, although straightforward, performs quite well on some of the experiments and is quite competitive with the first generation of gait recognition algorithms.

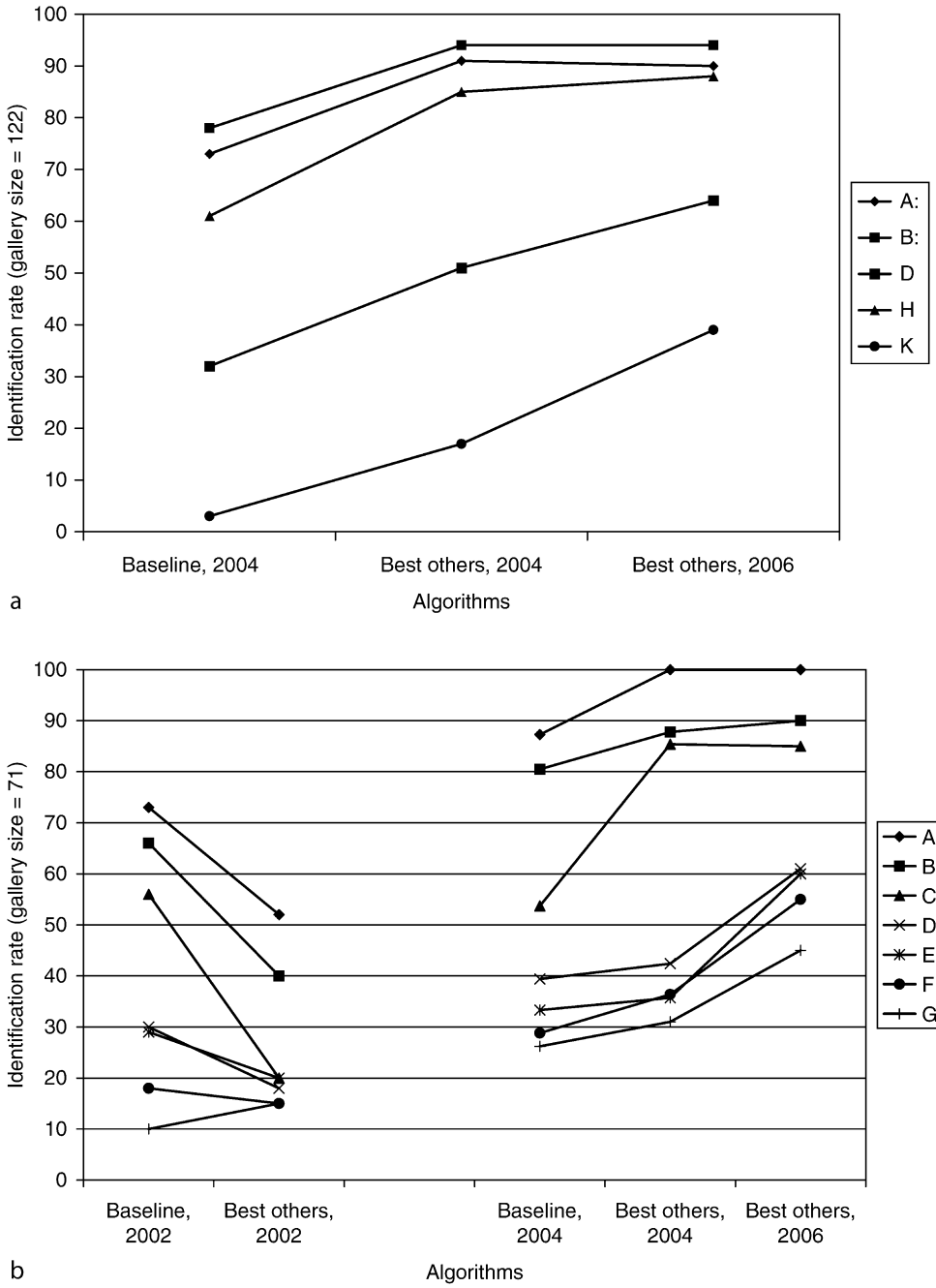
Performance on the Gait Challenge Problem

The results reported for the Gait Challenge problem are of two types, ones that report results on the first version of the dataset that was released with 71 subjects and the second set of results are those reported for the full dataset with 122 subjects. The smaller dataset allows just the first eight experiments listed in Table 1, but with reduced gallery set sizes. Figure 5a tracks the baseline performance and the best performance reported in the literature. As of middle of 2007, there were 18 papers that reported results on the smaller version of the problem. In 2002, when the Gait Challenge Problem was released, the performance of the baseline algorithm was better than the best reported performance. By 2004, while the baseline algorithm performance improved as the algorithm was fine-tuned, the performance of the best performance improved significantly and continued to improve through 2006. This trend is also seen in the results reported in six papers on the full dataset, summarized in Fig. 5b.

As is evident, the gait challenge problem has already spurred the development of gait recognition algorithms with improving performance. What is particularly interesting to notice is that the performance on hard experiments such those across surface (experiment D) and elapsed time (experiment K) has improved. Of course, there is still room for further improvement. Another interesting aspect is that the improvement of performance from 2004 to 2006 was not due to "continued engineering" of existing approaches, but involved the redesign of the recognition approaches. For instance, the greatest gains came from approaches that analyzed the silhouette shapes rather than the dynamics [2, 7]. Dynamics is important, but by itself is not sufficient.

Performance of a large number of algorithmic approaches have been explored. A review of the performances reported in these works reveals* [8] that

1. All most all of these approaches are based on silhouettes.
2. There is no one method that performs the best on all experiments.
3. Performances that involve matching against viewpoint and shoe variations, but on the same walking surface, has plateaued out.
4. Matching against walking surface variation remains a challenge.
5. Good performances ($>80\%$) has been reported for matching with and without carrying objects.
6. Matching across 6 months time difference has low performance, but the number of subjects involved in this experiment (33 subjects) is too low to derive meaningful conclusions.



Evaluation of Gait Recognition. Figure 5 Improvement in gait recognition algorithms over time with respect to the baseline performance. (**a**) Results on the first release of the gait dataset with 71 subjects in the gallery for the first eight experiments listed in Table 1 are tracked here. (**b**) Results on the full dataset with 122 subjects for the key experiments listed in Table 1. From 2004 to 2006, the best reported performances are better on all the experiments.

Other Large Datasets

There are currently two other datasets that are as large as the HumanID gait challenge dataset in terms of number of subjects. First is the CASIA Indoor Gait Database [4]. The gallery set includes 124 subjects with normal walking, no coat and no carry bag. Different probes can be defined in terms of changes in (1) viewpoint, (2) clothing change (coat vs. no-coat), and (3) carrying a bag and not carrying a bag. Not many algorithms have

reported performance on this dataset yet. But, the performance reported for the gait energy image approach in [9], seems to corroborate the findings from the HumanID Problem for matching across carrying conditions: a performance of upto 80% is reported for the CASIA dataset.

The other large dataset is the SOTON HID Gait Database [10] with 115 subjects are collected mostly indoor and some under outdoor conditions. The indoor SOTON dataset was collected to examine the premise that gait is unique so the background is controlled so as to allow easy segmentation. The same subjects were also filmed walking outdoors to determine whether gait biometrics could be perceived with complex backgrounds. Performances in the range of 72–85% have been reported for matching across sessions using a variety of approaches. This dataset also affords the matching across time issue. It has been shown that a time-dependent predictive model [11] results in 92% recognition, but only for ten subjects.

It is worth noting that face recognition on these data sets would be poor, indeed given the low-resolution and the uncontrolled lighting.

Future Evaluations

It is to be expected that each gait research group would collect their own data set to develop ideas. This is an important process. For instance, one new dataset is the CASIA infrared night gait dataset [12]. It consists of gait data from 153 subjects are collected outdoors, at night, with and without carrying condition, and at two different speeds. This dataset nicely complements existing datasets that are collected during the day. Given the data-driven nature of biometrics research, the key to future progress are such data sets, collected to explore issues not considered or raised by existing ones. For instance, as of today there is need for the better understanding of the variation of gait due to surface conditions and across elapsed time. Also, currently there is no dataset to explore the matching across time issue for a large number of subjects.

Ideally, the new datasets should consist of gait data from around 1,000 subjects, an order of magnitude larger than current large datasets. It is important to increase the number of subjects so, that it is possible to empirically study the scaling of performance with number of subjects. Some guidance about the required sizes can be found in [13, 14], where statistical reasoning is employed to relate the number of subjects with target error confidences. The data collection should include gait data repeated at regular time intervals of weeks, spanning about a year. The dataset should be collected in outdoor conditions, preferably collected at a distance of 300 m to reflect real world conditions. The dataset should come with a set of well defined experiments in terms of gallery and probe sets. These experiments will influence the types of algorithms. For the experiments to be effective at influencing the direction of gait research the design of the experiments needs to solve the *three bears problem*; the experiments must be neither too hard nor too easy, but just right. If performance on the experiments is easily saturated, then the gait recognition community will not be challenged. If experiments are too hard, then it will not be possible to make progress on gait recognition. Ideally, the set of experiments should vary in difficulty, characterize where the gait recognition problem is solvable, and explore the factors that affect performance. A set of experiments cannot meet this ideal unless the appropriate set of data is collected. It is important to view biometrics research as a data-driven algorithm development process rather than algorithm-driven data collection process.

Related Entries

- ▶ Hidden Markov Models
- ▶ Human Detection and Tracking
- ▶ Human Movement, Psychology
- ▶ Performance Evaluation, Overview
- ▶ Performance Testing Methodology Standardization
- ▶ Psychology of Gait Recognition
- ▶ Silhouette-Based Gait and Action Recognition
- ▶ Verification/Authentication/Identification/Recognition

References

1. Gross, R., Shi, J.: The CMU motion of body (MoBo) database. Tech. report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University (2001)

2. Kale, A., Sundaresan, A., Rajagopalan, A.N., Cuntoor, N.P., Roy-Chowdhury, A.K., Krüger, V., Chellappa, R.: Identification of humans using gait. IEEE Transactions on Image Processing **13**(9), 1163–1173 (2004)
3. Wagg, D.K., Nixon, M.S.: On automated model-based extraction and analysis of gait. In: International Conference on Automatic Face and Gesture Recognition, pp. 11–16 (2004)
4. Yu, S., Tan, D., Tan, T.: A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: International Conference on Pattern Recognition, vol. 4, pp. 441–444 (2006)
5. Sarkar, S., Jonathon Phillips, P., Liu, Z., Robledo-Vega, I., Grother, P., Bowyer, K.W.: The Human ID gait challenge problem: Data sets, performance, and analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(2), 162–177 (2005)
6. Liu, Z., Malave, L., Osuntogun, A., Sudhakar, P., Sarkar, S.: Toward understanding the limits of gait recognition. In: Proc. of SPIE Defense and Security Symposium: Biometric Technology for Human Identification, pp. 195–205 (2004)
7. Liu, Z., Sarkar, S.: Improved Gait Recognition by Gait Dynamics Normalization. IEEE Transactions on Pattern Analysis and Machine Intelligence **28**(6), 863–876 (2006)
8. Sarkar, S., Liu, Z.: Handbook of Biometrics, chap. Gait Recognition. Springer (2008)
9. Han, J., Bhanu, B.: Statistical feature fusion for gait-based human recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. II, pp. 842–847 (2004)
10. Shutler, J.D., Grant, M.G., Nixon, M.S., Carter, J.N.: On a large sequence-based human gait database. In: International Conference on Recent Advances in Soft Computing, pp. 66–71 (2002)
11. Veres, G., Nixon, M., Carter, J.: Model-based approaches for predicting gait changes over time. In: International Conference on Intelligent Sensors, Sensor Networks and Information, pp. 325–330 (2005)
12. Tan, D., Huang, K., Yu, S., Tan, T.: Efficient night gait recognition based on template matching. In: International Conference on Pattern Recognition, vol. 3, pp. 1000–1003 (2006)
13. Doddington, G., Przybocki, M., Martin, A., Reynolds, D.: The NIST speaker recognition evaluation–Overview, methodology, systems, results, perspective. Speech Communication **31**(2–3), 225–254 (2000)
14. Veres, G., Nixon, M., Carter, J.: Is Enough Enough?? What Is Sufficiency in Biometric Data?. Lecture Notes in Computer Science **4142**, 262 (2006)

Expected Performance or Utility of Fingerprint Image in an Automated Comparison Environment

► Fingerprint Image Quality

Expression

Face recognition processing usually requires a neutral facial expression (no smiling, mouth closed) and with the subject looking straight at the camera.

► Photography for Face Image Data

Extended Enterprise

Beyond the physical boundaries or "four walls" of an organization is extended enterprise.

► Remote Authentication

External Identification

Identification of a victim based on external evidence, such as the victim's gender, height, build, face, and fingerprints.

► Dental Biometrics

External Operation Time

► Operational Times

Extra-Class

Extra-class refers to instances of different subjects. Ideally, the extracted features should be very different for instances of different subjects.

► Local Image Filters

Eye Centers

In face recognition, eye center is defined as the geometric centroid of the close region formed by the upper and lower eyelids when the eye is opened. In practice, the midpoint of the left and right eye corners is often used instead. Because of the changing of gaze, pupil or origin of the iris circle is not necessarily the eye center.

► Face Misalignment Problem

Eye Tracking

The process of measuring either the motion of the eye relative to the head or the point of gaze, i.e. where someone is looking. Applications include medical and cognitive studies, computer interfaces, and marketing research.

► Segmentation of Off-Axis Iris Images

F

Face Acquisition

► Face Device

Face Aging

Face aging is to predict the future appearance of human face by learning the aging patterns, child growth, and adult aging are two type of aging.

► And-Or Graph Model for Faces

Face Alignment

Leon Gu, Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA

Synonyms

Face registration; Face matching

Definition

Face alignment is a computer vision technology for identifying the geometric structure of human faces in digital images. Given the location and size of a face, it automatically determines the shape of the face components such as eyes and nose. A face alignment program typically operates by iteratively adjusting a ► deformable models, which encodes the prior knowledge of face shape or appearance, to take into account the low-level image evidences and find the face that is present in the image.

Introduction

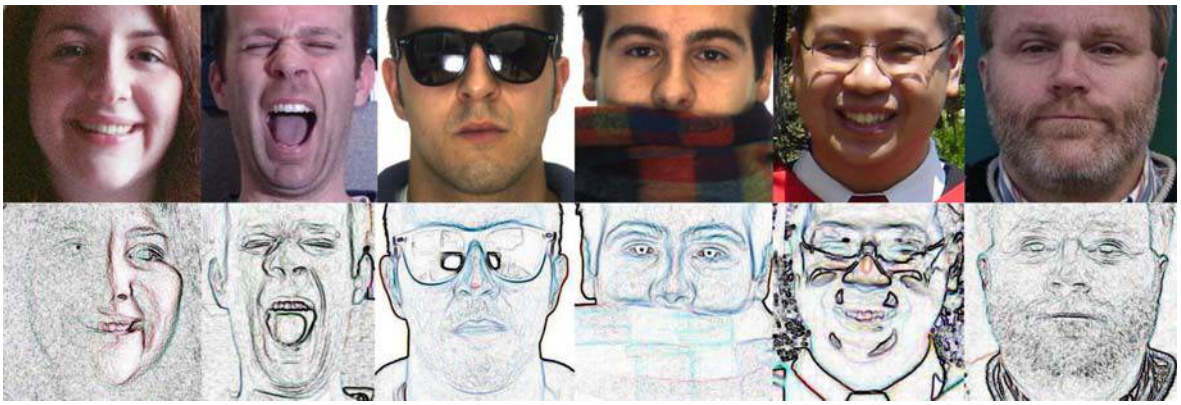
The ability of understanding and interpreting facial structures is important for many image analysis tasks. Suppose that, if we want to identify a person from a surveillance camera, a natural approach would be running the face image of the person through a database of known faces, examining the differences and identifying the best match. However, simply subtracting one image from another would not yield the desirable differences (as shown in Fig. 1), unless two faces are properly aligned. The goal of face alignment is to establish correspondence among different faces, so that the subsequent image analysis tasks can be performed on a common basis.

The main challenge in face alignment arises from pervasive ambiguities in low-level image features. Consider the examples shown in Fig. 2. While the main face structures are present in the ► feature maps, the contours of face components are frequently disrupted by gaps or corrupted by spurious fragments. Strong gradient responses could be due to reflectance, occlusion, fine facial texture, or background clutter. In contrast, the boundaries of face components such as nose and eyebrow are often obscure and incomplete. Looking for face components separately is difficult and often yields noisy results.

Rather than searching individual face components and expecting the face structure to emerge from the results, a better strategy is imposing the structure explicitly from the beginning. A majority of work in the field are developed based on this strategy. Deformable template [1], for example, is an elastic model which resembles face structure by assemblies of flexible curves. A set

© 2009 Springer Science+Business Media, LLC

Face Alignment. **Figure 1** To compare two face images, by directly adding them or subtracting one from another does not produce the desired result. Face alignment enables to establish correspondences between different images, so that the subsequent tasks can be performed on a common basis.



Face Alignment. **Figure 2** The major difficulty in face alignment is low-level image ambiguities. Face topologies could be significantly corrupted in the gradient feature maps (*second row*), due to various factors such as reflectance, occlusion, fine facial texture, and background clutter.

of model parameters control shape details such as the locations of various facial subparts and the angles of hinges which join them. The model is imposed upon and aligned to an image by varying the parameters. This strategy is powerful for resolving low-level image ambiguities. Inspired by this work, many variations of deformable face models emerged, including [2–9]. The common scheme in these work is first to construct a generic face model, then modify it to match the facial features found in a particular image. In this procedure, encoding prior knowledge of human faces, collecting image evidences of facial features, and fusing the observations with priors are the three key problems. Our treatment will follow the method proposed by Gu and Takeo [8, 9], which addresses the above problems in a coherent hierarchical Bayes framework.

Constructing Face Priors

This article concerns with the prior knowledge of a particular kind, namely shape priors. Suppose that, a face consists of a set of landmark points, which are typically placed along the boundaries of face components, i.e., $S = (x_1, y_1, \ldots, x_n, y_n)$. It can be viewed as a random vector, and its distribution, commonly called shape prior, describes the plausible spatial configurations of the landmark set. A principled way to construct the prior is by learning the distribution from training samples.

Face appears in different scales and orientations. First we need to transform all training face images into a common coordinate frame. One popular approach is general procrustes analysis [10]. It consists of two recursive steps: computing the mean shape, and

aligning each training shape with the mean by a rigid transformation. These two steps are repeated until the differences between the mean and the training shapes are minimized.

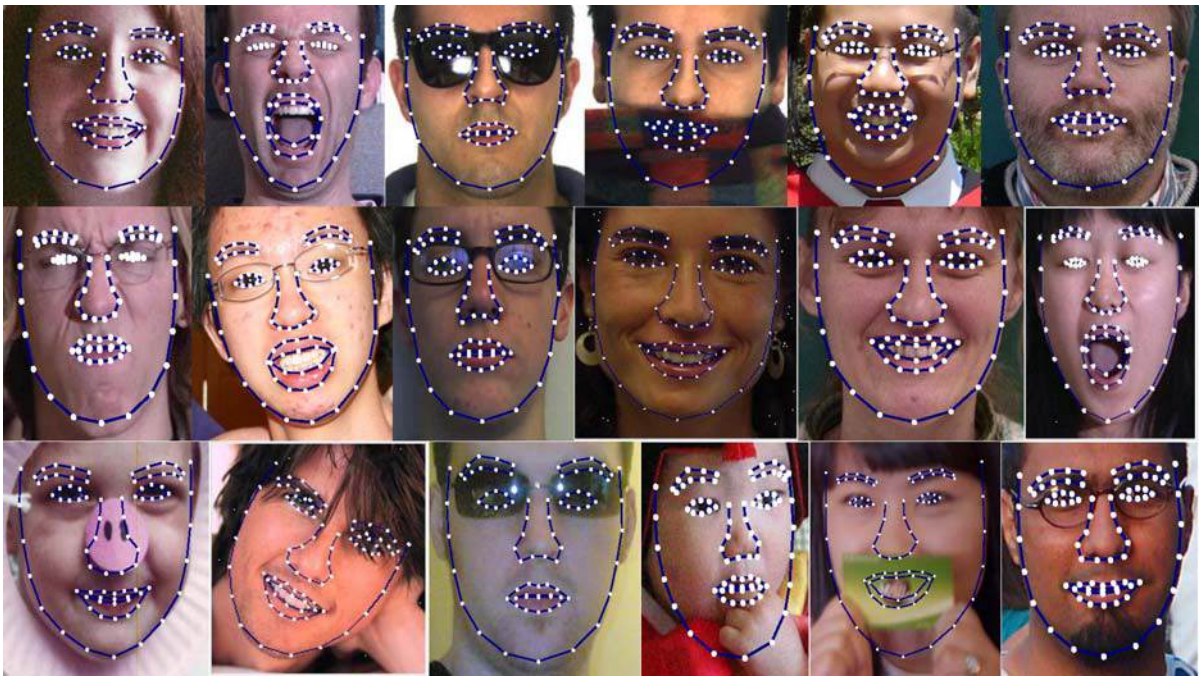
Next, we construct shape prior from the aligned training samples. The spatial arrangement of facial landmarks, although deformable, has to satisfy certain constrains. For example, it is often reasonable to assume that face shape is normally distributed, therefore, to learn the distribution we simply compute the mean and the covariance of the training shapes. More specifically, since the intrinsic variability of face structure is independent to its representation, e.g., the number of landmarks, we can parameterize face shape in a low-dimensional subspace [6, 8], such as

$$S = \Phi b + \mu + \epsilon. \quad (1)$$

The columns of Φ denote the major "modes" of shape deformations, and the elements of b controls the magnitude of deformation on the corresponding mode. This model has a nice generative interpretation: the shape vector S is generated by first adding a sequence of deformations $\{\Phi_i b_i\}$ into the mean shape μ, then permuting the resultant shape by an Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma^2)$. From a geometric perspective, the matrix Φ span a low-dimensional subspace which is centered at μ, the deformation coefficient b is the projection of S in the subspace, and ϵ denotes the deviation of S from the subspace. If assuming the elements of b to be independently normal, i.e., $b \sim \mathcal{N}(0, \Sigma)$ and Σ is diagonal, the distribution over the shape S is a constrained Gaussian, $S \sim \mathcal{N}(\mu, \Phi\Sigma\Phi^t + \sigma^2 I)$. The model parameters μ, Φ, Σ, and σ can be learned from training data. This model is also known as probabilistic principal component analysis [11] in the field of machine learning.

Detecting Facial Features

Strong gradient response is not the only way to characterize facial features. Some feature points may correspond to a weaker secondary edge in local context instead of the strongest; other points such as eye corners may have rich image structure that is more informative than gradient magnitude. Facial feature modeling can be made more effective by constructing detectors specific to each individual feature. One simple detector [2], for example, is a normal distribution built on the local gradient structures of each point. The distribution is learned



Face Alignment. Figure 3 Face alignment results from Gu and Kanade [9].

from training face images, and applied to evaluate the target image. Concatenating the best candidate position (u_i, v_i) of each feature point, we obtain an "observation" $Q = (u_1, v_1, \ldots, u_n, v_n)$ of the face shape that is likely to be present in the image. The observation is related with the aligned shape S by a rigid transformation

$$Q = \mathcal{T}(S, \theta) + \eta, \tag{2}$$

where $\theta = \{t, s, r\}$ denotes the transformation parameters (translation, scale, and rotation), and η is an additive observation noise. The conditional $p(Q|S)$ remains to be normal if the transformation $\mathcal{T}$ is linear, e.g., rigid or affine. More sophisticated detectors have been developed to produce better observations, however, after decades of research people have learned that individual feature detectors are effective only up to a point and cannot be expected to retrieve the entire face structure.

Fusing Prior with Image Observations

Combining the deformation model (1) with the transformation model (2) a hierarchical Bayes model is established that simulates how a random observation Q is generated from the deformation magnitude b and the transformation parameters θ. In this framework, the face alignment task is to modify shape priors to take into account the image evidences, arriving at the target face shape in images. EM algorithm is typically used for inferring the posterior b and θ, and analytic solutions exist for both E and M steps when the transformation is linear. This framework has been extended to model three-dimensional transformations for aligning multi-view faces [8], and nonlinear shape deformations for dealing with face images with exaggerated facial expressions [9]. Figure 3 shows a few alignment results from [9].

Summary

Significant progresses have been made in face alignment in recent years. The hierarchical Bayes formulation introduced in this article provides a systematic way to resolve low-level image ambiguities and exploit prior knowledge. Face alignment has a wide range of applications including face recognition, expression analysis, facial animation, lip reading, and human–computer interaction.

Related Entries

- ▶ Deformable Models
- ▶ Face Warping
- ▶ Feature Map

References

1. Yuille, A.L., Hallinan, P.W., Cohen, D.S.: Feature extraction from faces using deformable templates. Int. J. Comput. Vision **8**(2), 99–111 (1992). DOI http://dx.doi.org/10.1007/BF00127169. URL http://www.stat.ucla.edu/~yuille/pubs/optimize_papers/DT_IJCV1992.pdf
2. Cootes, T.F., Taylor, C., Cooper, D., Graham, J.: Active shape models – their training and their applications. Comput. Vision Image Understanding (1995)
3. Wiskott, L., Fellous, J.M., Kruger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 775–779 (1997). DOI http://dx.doi.org/10.1109/34.598235. URL http://www.face-rec.org/algo rithms/EBGM/WisFelKrue99-FaceRecognition-JainBook.pdf
4. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3d-faces. In: ACM SIGGRAPH (1999)
5. Cootes, T., Edwards, G., Taylor, C.: Active appearance models **23**(6), 681–685 (2001)
6. Zhou, Y., Gu, L., Zhang, H.: Bayesian tangent shape model: Estimating shape and pose parameters via Bayesian inference, pp. I: 109–116 (2003). URL http://www.cs.cmu.edu/~gu/publication/alignment_cvpr03.pdf
7. Zhang, Z., Liu, Z., Adler, D., Cohen, M.F., Hanson, E., Shan, Y.: Robust and rapid generation of animated faces from video images – a model-based modeling approach. Int. J. Comput. Vision (2004)
8. Gu, L., Kanande, T.: 3d alignment of face in a single image. In: CVPR (2006)
9. Gu, L., Kanade, T.: A generative shape regularization model for robust face alignment. In: The Tenth European Conference on Computer Vision (2008)
10. Goodall, C.: Procrustes methods in the statistical analysis of shape. J. Royal Statistical Society. Series B (Methodological) **53**, 285–339 (1991).
11. Jipping, M., Bishop, C: Probabilistic principal component analysis. J. Royal Statistical Society (1999).

Face Alignment Error

- ▶ Face Misalignment Problem

Face Biometric

▶ Face Recognition, Overview

Face Camera

▶ Face Device

Face Databases and Evaluation

Dmitry O. Gorodnichy
Laboratory and Scientific Services Directorate, Canada Border Services Agency, Ottawa, ON, Canada

Synonyms

Face recognition performance evaluation

Definition

Face Databases are imagery data that are used for testing ▶ face processing algorithms. In the contents of biometrics, face databases are collected and used to evaluate the performance of face recognition biometric systems.

Face recognition evaluation is the procedure that is used to access the recognition quality of a face recognition system. It involves testing the system on a set of face databases and/or in a specific setup for the purpose of obtaining measurable statistics that can be used to compare systems to one another.

Introduction: Factors Affecting Face Recognition Performance

While for humans recognizing a face in a photograph or in video is natural and easy, computerized face recognition is very challenging. In fact, automated recognition of faces is known to be more difficult than recognition of other imagery data such as iris, vein, or fingerprint images due to the fact that the human face is a non-rigid 3D object which can be observed at different angles and which may also be partially occluded. Specifically, face recognition systems have to be evaluated with respect to the following factors [1]:

1. Face image resolution – face images can be captured at different resolutions: face images scanned from documents may have very high resolution, while face captured with a video camera will mostly be of very low resolution,
2. Facial image quality – face images can be blurred due to motion, out of focus, and of low contrast due to insufficient camera exposure or aperture, especially when captured in uncontrolled environment,
3. Head orientation – unless a person is forced to face the camera and look straight into it, will unlikely be seen under the same orientation on the captured image,
4. Facial expression – unless a person is quiet and motionless, the human face constantly exhibits a variety of facial expressions
5. Lighting conditions – depending on the location of the source of light with respect to the camera and the captured face, facial image will seen with different illumination pattern overlaid on top of the image of the face,
6. Occlusion – image of the face may be occluded by hair, eye-glasses and clothes such scarf or handkerchief,
7. Aging and facial surgery – compared to fingerprint or iris, person faces changes much more rapidly with time, it can also be changed as a result of make-up or surgery.

There are over thirty publicly available face databases. In addition, there are Face Recognition Vendor Test (FRVT) databases that are used for independent evaluation of Face Recognition Biometric Systems (FRBS). Table 1 summarizes the features of the most frequently used still image facial databases, as pertaining to the performance factors listed above. More details about each database can be found at [2–4] and below are presented some of them. For the list of some video-based facial databases, see [5].

Public Databases

One of the first and most used databases is AT&T (formerly "Olivetti ORL") database [6] that contains

ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open/closed eyes, smiling/not smiling) and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).

The other most frequently used dataset is developed for FERET program [7]. The images were collected in a semi-controlled environment. To maintain a degree of consistency throughout the database, the same physical setup was used in each photography session. A duplicate set is a second set of images of a person already in the database and was usually taken on a different day. For some individuals, over 2 years had elapsed between their first and last sittings, with some subjects being photographed multiple times.

The Yale Face Database [8] contains images of different facial expression and configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink. The Yale Face Database B provides single light source images of 10 subjects each seen under 576 viewing conditions (9 poses x 64 illumination conditions). For every subject in a particular pose, an image with ambient (background) illumination was also captured.

The BANCA multi-modal database was collected as part of the European BANCA project, which aimed at developing and implementing a secure system with enhanced identification, authentication, and access control schemes for applications over the Internet [9]. The database was designed to test multimodal identity verification with various acquisition devices (high and low quality cameras and microphones) and under several scenarios (controlled, degraded, and adverse).

To investigate the time dependence in face recognition, a large database is collected at the University of Notre Dame [10]. In addition to the studio recordings, two images with unstructured lighting are obtained.

University of Texas presents a collection of a large database of static digital images and video clips of faces [11]. Data were collected in four different categories: still facial mug shots, dynamic facial mug shots, dynamic facial speech and dynamic facial expression. For the still facial mug shots, nine views of the subject, ranging from left to right profile in equal-degree steps were recorded. The sequence length is cropped to be 10 s.

The AR Face Database [12] is one of the largest datasets showing faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf).

XM2VTS Multimodal Face Database provides five shots for each person [13]. These shots were taken at one week intervals or when drastic face changes occurred in the meantime. During each shot, people have been asked to count from "0" to "9" in their native language (most of the people are French speaking), rotate the head from 0 to −90 degrees, again to 0, then to +90 and back to 0 degrees. Also, they have been asked to rotate the head once again without glasses if they wear any.

CMU PIE Database is one of the largest datasets contains images of 68 people, each under 13 different poses, 43 different illumination conditions, and with four different expressions [14].

The Korean Face Database (KFDB) contains facial imagery of a large number of Korean subjects collected under carefully controlled conditions [15] . Similar to the CMU PIE database, this database has images with varying pose, illumination, and facial expressions were recorded. In total, 52 images were obtained per subject. The database also contains extensive ground truth information. The location of 26 feature points (if visible) is available for each face image.

CAS-PEAL Face Database is another large-scale Chinese face database with different sources of variations, especially Pose, Expression, Accessories, and Lighting [16].

FRVT Databases

Face Recognition Vendor Tests (FRVT) provide independent government evaluations of commercially available and prototype face recognition technologies [4]. These evaluations are designed to provide U.S. Government and law enforcement agencies with information to assist them in determining where and how facial recognition technology can best be deployed. In addition, FRVT results serve to identify future research directions for the face recognition community. FRVT 2006 follows five previous face recognition technology evaluations – three FERET evaluations (1994, 1995 and 1996) and FRVT 2000 and 2002.

FRVT provides two new datasets that can be used for the purpose: high computational intensity test (HCInt) data set and Medium Computational Intensity test (MCInt) data set. HCInt has 121,589 operational well-posed (i.e. frontal to within 10 degrees) images of 37,437 people, with at least three images of each person.

The images are provided from the U.S. Department of State.s Mexican non-immigrant visa archive. The images are of good quality and are gathered in a consistent manner, collected at U.S. consular offices using standard issue digital imaging apparatus whose specification remained fixed over the collection period.

The MCInt data set is composed of a heterogeneous set of still images and video sequences of subjects in a variety of poses, activities and illumination conditions. The data are collected from several sources, captured indoors and outdoors, and include lose-range video clips and static images (with over hundred individuals), high quality still images, Exploration Video Sequences (where faces move through the nine facial poses used for the still images) and Facial Speech Videos (where two video clips were taken of individuals speaking, first in a neutral way, then in an animated way).

Face Evaluation

For an evaluation to be accepted by the biometric community, the performance results have to published along with the evaluation protocol. An evaluation protocol describes how the experiments are run and how the data are collected. It should be written in sufficient detail so that users, developers, and vendors can repeat the evaluation.

The main attributes of the evaluation protocol are described below.

Image Domain and Face Processing Tasks

There are two image domains where Face Recognition Biometric Systems (FRBS) are applied:

1. *Face recognition in documents* (FRiD), in particular, face recognition from Machine Readable Travel Documents (MRTD).
2. *Face recognition in video* (FRiV), also referred to as *Face in Crowd* problem, an example of which is face recognition from surveillance video and TV.

These two image domains are very different [17]. The systems that perform well in one domain may not perform well in the other [18].

FRiD deals with facial data that are of high spacial resolution, but that are very limited or absent in ▶ temporal domain – FRiD face images would normally have *intra-ocular distance* (IOD) of at least 60 pixels, which is the distance defined by the ▶ canonical face model established by International Civil Aviation Organization (ICAO) for MRTD. There will however be not more than one or very few images available of the same person captured over a period of time.

In contrast, FRiV deals with facial images that are available in abundance in temporal domain but which are of much lower spatial resolution. The IOD of facial images in video is often lower than 60 pixels, due to the fact that face normally occupies less than one eighth of a video image, which itself is relatively small (352×240 for analog video or 720×480 for digital video) compared to a scanned document image. In fact, IOD of faces detected in video is often just slightly higher than or equal to 10 pixels, which is the minimal IOD that permits automatic detection of faces in images [19].

While for FRiD facial images are often extracted beforehand and face recognition problem is considered in isolation from other face processing problems, FRiV requires that a system be capable of performing several other facial processing tasks prior to face recognition, such as face detection, face tracking, eye localization, best facial image selection or reconstruction, which may also be coupled with facial image accumulation and video snapshot resolution enhancement [20]. Evaluation of FRBS for FRiD is normally performed by testing a system on static facial images datasets described above. To evaluate FRBS for FRiV however, it is much more common to see the system testing performed as a pilot project on a real-life video monitoring surveillance task [21], although some effort to evaluate their performance using prerecorded datasets and motion pictures has been also suggested and performed [5].

Use of Color

Color information does not affect the face recognition performance [22], which is why many countries still allow black-n-white face pictures in passport documents. Color however plays an important role in face detection and tracking as well as in eye localization. Therefore, for testing recognition from video, color video streams should be used.

Scenario Taxonomy

The following scenario taxonomy is established to categorize the performance of biometric systems [23]:

Database (year created)	#individuals/ # images	i.o.d/ image width	Orientation	Expression	Lighting /quality	Occlusion	Situations	Representative Facial image
AT &T Olivetti (1992–1994)	40/400	~60/92	yes	yes	yes	yes	yes	
FERET (1993–1996)	1999/ 14,126	~80/256	9–20	2	2		2	
AR	116/ 3288	~90/768	1	4	4	2	2	
Yale (B)	15/165 10/5760	~80/640	9		64			
PIE 2000	68/ 41,368	~75/640	13	3	43			
Korean	1000/ 52000	~80/640	7	5	16			
Cas-peal 2003	1040	~45/360	21	15	6		1–5	
Human ID	350/ 15.500	~80/1600	1	2	3		10	
UofT 2002	284	~80/720	video	video				
Banca 2002–2003	208/ 208*12	~45/720	1	yes	3		12	
XM2VTS	293/	~100/720	Full rotation	speaking		Yes - eyeglasses	4	
Equinox	91	~100/240	1	3	3			
Cmu-hyperspectral	54	80/640	1		4		5	
nist	573/3248	~80	2					
FRVT HCInt	37,437/ 121,589	1					3	
FRVT MCInt 1999–2002			several	Still and video	several			

Face Databases and Evaluation. Figure 1 Face databases, categorized by the factors affecting the performance of face recognition systems: such as number of probes, face image resolution, head orientation, face expression, changed in lighting, image quality degradation, occlusion, and aging.

cooperative vs. non-cooperative, overt vs. covert, habituated vs. non-habituated, attended vs. non-attended, public vs. private, standard vs. non-standard. When performing evaluation of FRBS, these categories have to be indicated.

Dataset Type and Recognition Task

Two types of datasets are possible for recognition problems:

1. Closed dataset, where each query face is present in the database, as in a watch list in the case of negative enrollment, or as in a list of computer users or ATM clients, in the case of positive enrollment,
2. Open dataset, where query faces may not be (or very likely are not) in the database, as in the case of surveillance video monitoring.

FRBS can be used for one three face recognition tasks:

1. Face verification, also referred to as authentification or 1 to 1 recognition, or positive verification, as verifying ATM clients,
2. Face identification, also referred to as or 1 to N (negative identification – as detecting suspects from a watch list), where a query face is compared against all faces in a database and the best match (or the best k matches) are selected to identify a person.
3. Face classification, also referred to as categorization, where a person is recognized as belonging to one of the limited number of classes, such as describing the person's gender (male, female), race (caucasian, asian etc), and various medical or genetic conditions (Down's Syndrome etc).

While the result of the verification and identification task are used as hard biometrics, the results from classification can be used as *soft biometrics*, similar to person's height or weight.

Performance Measures

The performance is evaluated against two main errors a system can exhibit:

1. False accept (FA) also known as false match (FM), false positive (FP) or type I error.
2. False reject (FR) also known as false non-match (FNM) or false negative (FN) or type 2 error.

By applying a FRBS on a significantly large data set of facial images, the total number of FA and FR are measured and used to compute one or several of the following cumulative measurements and figures of merit (FOM). For verification systems,

1. FA rate (FAR) with fixed FR rate.
2. FR rate (FRR), or true acceptance rate (TAR = 1 – FRR), also known as true positive (or hit) rate, at fixed FA rate.
3. Detection Error Trade-off (DET) curve, which is the graph of FAR vs FRR, which is obtained by varying the system parameters such as *match threshold*.
4. Receiver Operator Characteristic (ROC) curve, which is similar to DET curve, but plots TAR against FAR.
5. Equal error rate (EER), which the FAR measured when it equals FRR.

For identification tasks,

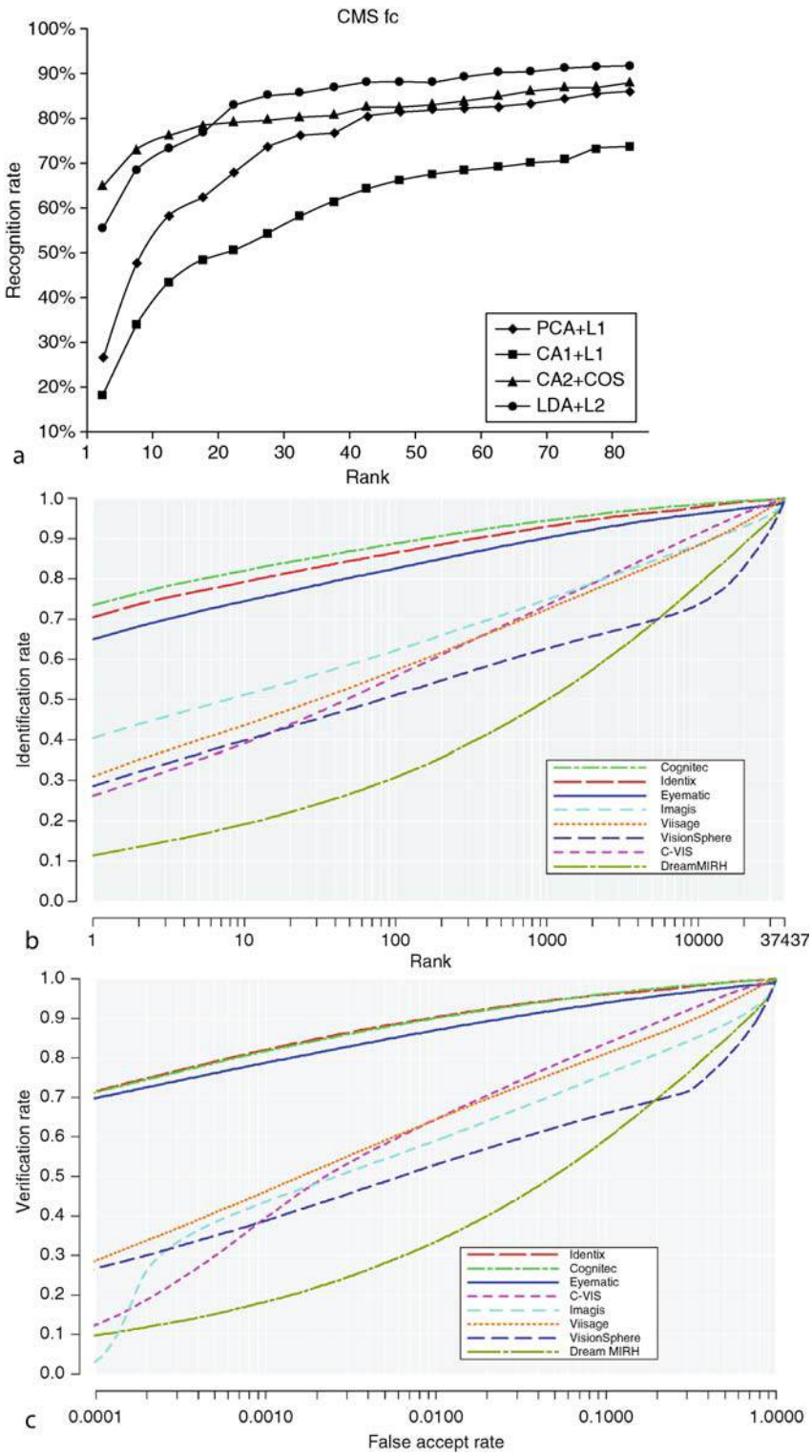
1. Identification rate, or rank-1 identification, which is number of times when the correct identity is chosen as the most likely candidate.
2. Rank-k identification rate (Rk), which is number of times the the correct identity is in the top k most likely candidates.
3. Cumulative Match Characteristic (CMC), which plots the rank-k identification rate against k.

The rates are counted as percentages to the number of faces in a databases. DET and ROC curves are often plotted using logarithmic axes to better differentiate the systems that shows similar performance.

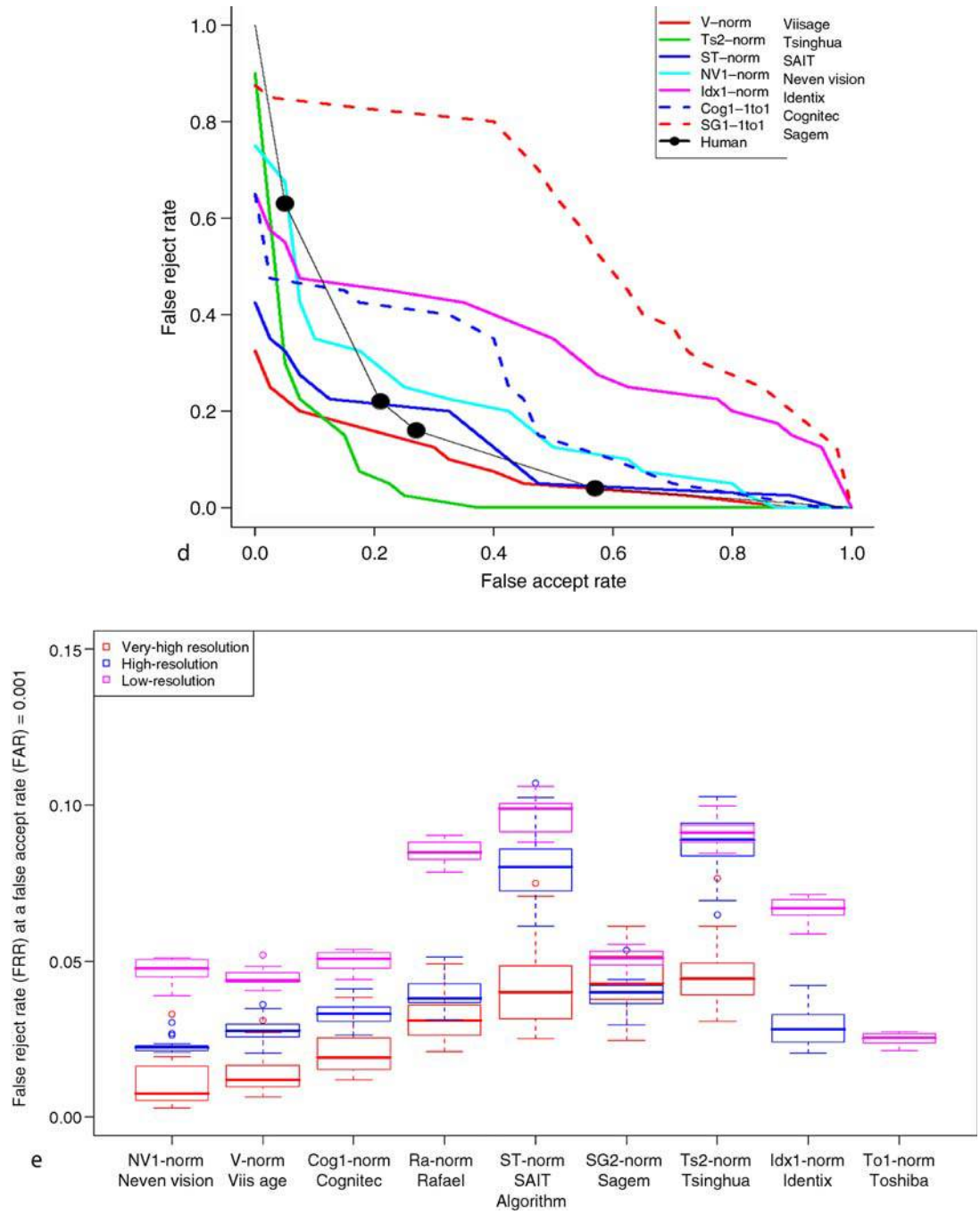
Similarity Metrics, Normalization, and Data Fusion

Different types of metrics can be used to compare ▶ feature vectors of different faces to one another. The recognition results can also be normalized. Proper covariance-weighted metrics and normalization should be used when comparing the performance results obtained on different datasets.

When temporal data are available, as when recognizing a person from a video sequence, the recognition results are often integrated over time in a procedure known as evidence accumulation or data fusion. The details of this should be known.



Face Databases and Evaluation. Figure 2 (Continued)



Face Databases and Evaluation. Figure 2 Examples of performance evaluation conducted on face databases: (**a**) identification performance of several appearance-based recognition algorithms measured using CMC curves on FERET database (from [25]), (**b–e**) verification and identification performance of commercial face recognition biometrics systems on FRVT datasets (from [24, 26], using CMC curves (**b**), ROC curve (**c**), DET curve (**d**) and fixed-FAR FRR distrubutions (**e**)).

Example Protocols

Feret protocol [7] is an example of the close set face identification, where a full distance matrix that measures the similarity between each query image and each database image is computed. FRVT2002 [24] addresses both open set verification problem and close-set identification problem and uses CMC and ROC to compare the results. BANCA protocol [9], which is designed for multi-modal databases, is an example of the open set verification protocol. XM2VTS Lausanne protocol [13] is an example of a close set verification, where anyone not in the database is considered an imposter.

Evaluation Results

Face Databases have been used over the years to compare and improve the existing face recognition techniques. Some of the obtained evaluation results are shown in Fig. 2. Figure 2a shows face identification results from [25] for popular appearance-based face-recognition techniques: Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Linear Discriminant Analysis (LDA), obtained on FERET database using CMC curves.

Figures 2b–e show performance evaluation of commercial FRBSs that participated in the FRVT2002 and FRVT2006 tests taken from [24, 26]

Future Work

Considerable advances have been made recently in the area of automated face recognition. FRBSs are now able to *recognize faces in documents* with the performance that matches or exceeds the human recognition performance. In large part, this has become possible due to the help of many researchers that have collected and maintained face databases. At the same time, despite the intensive use of these databases, no FRBS has been developed so far that can *recognize faces from video* with performance close to that of humans.

Automated recognition of faces from video is considerably worse than face recognition from documents, whereas for humans it is known to be the opposite. This status-quo situation serves as an indication that new evaluation datasets and benchmarks are needed for testing video-based face recognition systems. Knowing how easily available have become recently amounts of various video data (including news casts, televised shows, motion pictures, etc), it is foreseen that instead of using video-based data-bases, which are very costly and time consuming to create, the research community will soon adopt face evaluation benchmarks and protocols based on public domain video recordings [5].

The importance of improving the performance of video-based face recognition should not be underestimated, taking into account that of all hard biometric modalities, video-based face recognition is the most collectable and acceptable [27].

Related Entries

- ▶ Face Detection
- ▶ Face Recognition
- ▶ Identification
- ▶ Verification

References

1. Gorodnichy, D.O.: Facial recognition in video. In: Proceedings of the International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA'03), LNCS 2688, pp. 505–514, Guildford, UK, online at http://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-47150.pdf. (2003)
2. Gross, R.: Face Databases. Springer, New York (2005)
3. Face Recognition website. http://www.face-rec.org
4. Face Recognition Vendor Test website. http://www.frvt.org
5. Gorodnichy, D.O.: Seeing faces in video by computers (Editorial). Image and Video Computing, Special Issue on Face Processing in Video Sequences (online at http://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-48295.pdf) **24,** 1–6 (2006)
6. AT&T.: The Database of Faces (formerly The ORL Database of Faces, http://www.cl.cam.ac.uk/research/dtg/attarchive/facesataglance.html
7. Phillips, P.J., Moon, H., Rizvi, S., Rauss, P.J.: The FERET evaluation methodology for face-recognition algorithms. IEEE Trans. Pattern Anal. Mach. Intell. **22**(10), 1090–1104 (2000). http://www.nist.gov/humanid/feret/
8. Georghiades, A., Kriegman, D., Belhumeur, P.: From few to many: generative models for recognition under variable pose and illumination. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 643–660 (2001)
9. Bailly-Bailliere, E., Bengio, S., Bimbot, F., Hamouz, M., Kittler, J., Mariethoz, J., Matas, J., Messer, K., Popovici, V., Poree, F., Ruiz, B., Thiran, J.-P.: The BANCA database and evaluation protocol. In Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 625–638 (2003)

10. Phillips, P.J.: Human identification technical challenges. In IEEE International Conference on Image Processing, vol. 1, pp. 22–25 (2002)
11. OToole, A., Harms, J., Snow, S., Hurst, D.R. Pappas, M., Abdi, H.: A video database of moving faces and people, submitted (2003)
12. Martinez, A.M., Benavente, R.: The AR face database. Technical Report 24, Computer Vision Center(CVC) Technical Report, Barcelona (1998)
13. Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G.: XM2VTSDB: The extended M2VTS database. In: Second International Conference on Audio and Video-based Biometric Person Authentication (1999)
14. Sim, T., Baker, S., Bsat, M.: The CMU pose, illumination, and expression database. IEEE Trans. Pattern Anal. Mach. Intell. **25**(12), 1615–1618, http://www.ri.cmu.edu/projects/project 418.html) (2003)
15. Hwang, B.-W., Byun, H., Roh, M.-C., Lee, S.-W.: Performance evaluation of face recognition algorithms on the asian face database, KFDB. In Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 557–565 (2003)
16. Gao, W., Cao, B., Shan, S., Zhou, D., Zhang, X. Zhao, D.: CAS-PEAL large-scale Chinese face database and evaluation protocols. Technical Report JDL-TR-04-FR-001, Joint Research and Development Laboratory, http://www.jdl.ac.cn/peal (2004)
17. Gorodnichy, D.O.: Video-based framework for face recognition in video. In: Second International Workshop on Face Processing in Video (FPiV'05). Proceedings of Second Canadian Conference on Computer and Robot Vision (CRV'05), pp. 330–338. Victoria, BC, Canada, ISBN 0-7695-2319-6, online at http://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-48216.pdf. (2005)
18. Gorodnichy, D.O.: Recognizing faces in video requires approaches different from those developed for face recognition in photographs. In: Proceedings of NATO IST - 044 Workshop on Enhancing Information Systems Security through Biometrics. Ottawa, ON, Canada, October 18–20 (online at http://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-47149.pdf. (2004)
19. Shakhnarovich, G., Viola, P.A., Moghaddam, B.: A unified learning framework for realtime face detection and classification. In: International Conference on Automatic Face and Gesture Recognition, pp. 10–15, USA (2002)
20. Gorodnichy, D.O.: Introduction to the First IEEE Workshop on Face Processing in Video. In: First IEEE CVPR Workshop on Face Processing in Video (FPIV'04), Washington DC, USA, online at http://www.visioninterface.net/fpiv04/preface.html (2004)
21. Willing, R.: Airport anti-terror systems flub tests face-recognition technology fails to flag suspects, in USA TODAY. Accessed Sept 4, 2003. http://www.usatoday.com/usatonline/20030902/5460651s.htm.
22. Yip, A., Sinha, P.: Role of color in face recognition. MIT tech report (ai.mit.com) AIM-2001-035 CBCL-212 (2001)
23. Wayman, J.L., Jain, A.K., Maltoni, D., Maio, D. (eds.): Biometric Systems: Technology, Design and Performance Evaluation. Springer, New York (2005)
24. Phillips, P.J., Grother, P., Ross, J.M., Blackburn, D., Tabassi, E., Bone, M.: Face recognition vendor test 2002: evaluation report (March 2003)
25. Delac, K., Grgic, M., Grgic, S.: Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set. Int. J. Imaging Syst. Technol. **15**(5), 252–260 (2006)
26. Overview of the Face Recognition Grand Challenge - IEEE Conference on Computer Vision and Pattern Recognition, June 2005. Online at http://www.frvt.org/FRGC
27. Jain, A.K., Ross, A., Prabhakar, S.: An Introduction to Biometric Recognition. IEEE Trans. Circ. Syst. Video Technol. Special Issue on Image- and Video-Based Biometrics **14**, 4–20 (2004)

Face Detection

Ming-Hsuan Yang
University of California, Merced, CA, USA

Synonym

Face Localization

Definition

Face detection is concerned with finding whether there are any faces in a given image (usually in gray scale) and, if present, return the image location and content of each face. This is the first step of any fully automatic system that analyzes the information contained in faces (e.g., identity, gender, expression, age, race, and pose). While earlier work dealt mainly with upright frontal faces, several systems have been developed that are able to detect faces fairly accurately with in-plane or out-of-plane rotations in real time. Although a face detection module is typically designed to deal with single images, its performance can be further improved if video stream is available.

Introduction

The advances of computing technology has facilitated the development of real-time vision modules that interact with humans in recent years. Examples abound, particularly in biometrics and human computer interaction as the information contained in faces needs to be analyzed for systems to react accordingly. For biometric systems that use faces as nonintrusive input modules,

it is imperative to locate faces in a scene before any recognition algorithm can be applied. An intelligent vision-based user interface should be able to tell the attention focus of the user (i.e., where the user is looking at) in order to respond accordingly. To detect facial features accurately for applications such as digital cosmetics, faces need to be located and registered first to facilitate further processing. It is evident that face detection plays an important and critical role for the success of any face processing systems.

The face detection problem is challenging as it needs to account for all possible appearance variation caused by change in illumination, facial features, occlusions, etc. In addition, it has to detect faces that appear at different scale, pose, with in-plane rotations. In spite of all these difficulties, tremendous progress has been made in the last decade and many systems have shown impressive real-time performance. The recent advances of these algorithms have also made significant contributions in detecting other objects such as humans/pedestrians, and cars.

Operation of a Face Detection System

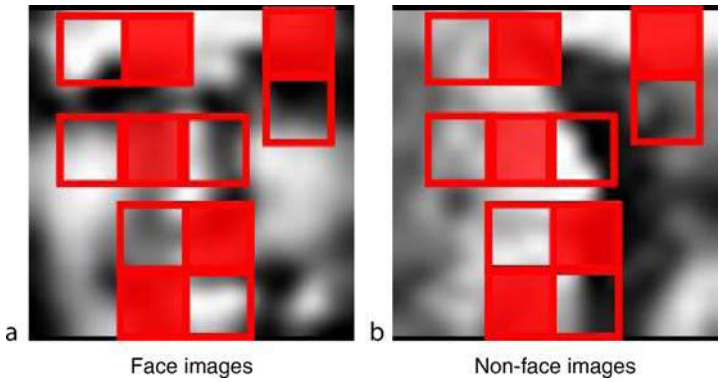
Most detection systems carry out the task by extracting certain properties (e.g., local features or holistic intensity patterns) of a set of training images acquired at a fixed pose (e.g., upright frontal pose) in an off-line setting. To reduce the effects of illumination change, these images are processed with histogram equalization [1, 2] or standardization (i.e., zero mean unit variance) [3]. Based on the extracted properties, these systems typically scan through the entire image at every possible location and scale in order to locate faces. The extracted properties can be either manually coded (with human knowledge) or learned from a set of data as adopted in the recent systems that have demonstrated impressive results [1, 2, 3, 4, 5]. In order to detect faces at different scale, the detection process is usually repeated to a pyramid of images whose resolution are reduced by a certain factor (e.g., 1.2) from the original one [1, 2]. Such procedures may be expedited when other visual cues can be accurately incorporated (e.g., color and motion) as pre-processing steps to reduce the search space [5]. As faces are often detected across scale, the raw detected faces are usually further processed to combine overlapped results and remove false positives with heuristics (e.g., faces typically do not overlap in images) [2] or further processing (e.g., edge detection and intensity variance).

Numerous representations have been proposed for face detection, including pixel-based [1, 2, 5], parts-based [4, 6, 7], local edge features [8], Haar wavelets [4, 9], and ▶ Haar-like features [3, 10]. While earlier holistic representation schemes are able to detect faces [1, 2, 5], the recent systems with Haar-like features [3, 11, 12] have demonstrated impressive empirical results in detecting faces under occlusion. A large and representative training set of face images is essential for the success of learning-based face detectors. From the set of collected data, more positive examples can be synthetically generated by perturbing, mirroring, rotating, and scaling the original face images [1, 2]. On the other hand, it is relatively easier to collect negative examples by randomly sampling images without face images [1, 2].

As face detection can be mainly formulated as a pattern recognition problem, numerous algorithms have been proposed to learn their generic templates (e.g., eigenface and statistical distribution) or discriminant classifiers (e.g., neural networks, Fisher linear discriminant, sparse network of Winnows, decision tree, Bayes classifiers, support vector machines, and ▶ AdaBoost). Typically, a good face detection system needs to be trained with several iterations. One common method to further improve the system is to bootstrap a trained face detector with test sets, and retrain the system with the false positive as well as negatives [2]. This process is repeated several times to further improve the performance of a face detector. A survey on these topics can be found in [5], and the most recent advances are discussed in the next section.

Recent Advances

The AdaBoost-based face detector by Viola and Jones [3] demonstrated that faces can be fairly reliably detected in real-time (i.e., more than 15 frames per second on 320 × 240 images with desktop computers) under partial occlusion. While Haar wavelets were used in [9] for representing faces and pedestrians, they proposed the use of Haar-like features which can be computed efficiently with integral image [3]. Figure 1 shows four types of Haar-like features that are used to encode the horizontal, vertical, and diagonal



Face Detection. **Figure 1** Four types of Haar-like features. These features appear at different position and scale. The Haar-like features are computed as the difference of dark and light regions. They can be considered as features that collect local edge information at different orientation and scale. The set of Haar-like features is large, and only a small amount of them are learned from positive and negative examples for face detection.

intensity information of face images at different position and scale.

Given a sample image of 24 × 24 pixels, the exhaustive set of parameterized Haar-like features (at different position and scale) is very large (about 160,000). Contrary to most of the prior algorithms that use one single strong classifier (e.g., neural networks and support vector machines), they used an ensemble of weak classifiers where each one is constructed by thresholding of one Haar-like feature. The weak classifiers are selected and weighted using the AdaBoost algorithm [13]. It is worth to note that boosting algorithms can also be derived from the perspective of function approximation with gradient descent and applications for regression [14]. As there are large number of weak classifiers, they presented a method to rank these classifiers into several cascades using a set of optimization criteria. Within each stage, an ensemble of several weak classifiers is trained using the AdaBoost algorithm. The motivation behind the cascade of classifier is that simple classifiers at early stage can filter out most negative examples efficiently, and stronger classifiers at later stage are only necessary to deal with instances that look like faces. The final detector, a 38 layer cascade of classifiers with 6,060 Haar-like features, demonstrated impressive real-time performance with fairly high detection and low false positive rates. Several extensions to detect faces in multiple views with in-plane ration have since been proposed [11, 12, 15]. An implementation of the AdaBoost-based face detector [3] can be found in the Intel OpenCV library.

Despite the excellent run-time performance of boosted cascade classifier [3], the training time of such a system is rather lengthy. In addition, the ▶ classifier cascade is an example of degenerate decision tree with an unbalanced data set (i.e., a small set of positive examples and a huge set of negative ones). Numerous algorithms have been proposed to address these issues and extended to detect faces in multiple views. To handle the asymmetry between the positive and negative data sets, Viola and Jones proposed the asymmetric AdaBoost algorithm [16] which keeps most of the weights on the positive examples. In [3], the AdaBoost algorithm is used to select a specified number of weak classifiers with lowest error rates for each cascade and the process is repeated until a set of optimization criteria (i.e., the number of stages, the number of features of each stage, and the detection/false positive rates) is satisfied. As each weak classifier is made of one single Haar-like feature, the process within each stage can be considered as a feature selection problem. Instead of repeating the feature selection process at each stage, Wu et al. [17] presented a greedy algorithm for determining the set of features for all stages first before training the cascade classifier. With the greedy feature selection algorithm used as a pre-computing procedure, they reported that the training time of the classifier cascade with AdaBoost is reduced by 50–100 times. For learning in each stage (or node) within the classifier cascade, they also exploited the asymmetry between positive and negative data using a linear classifier with the assumption that they can be modeled with Gaussian distributions [17]. The merits and drawbacks of the

proposed linear asymmetric classifier as well as the classic Fisher linear discriminant were also examined in their work. Recently, Pham and Cham proposed an online algorithm that learns asymmetric boosted classifiers [18] with significant gain in training time.

In [19], an algorithm that aims to automatically determine the number of classifiers and stages for constructing a boosted ensemble was proposed. While a greedy optimization algorithm was employed in [3], Brubaker et al. proposed an algorithm for determining the number of weak classifiers and training each node classifier of a cascade by selecting operating points within a receiver operator characteristic (ROC) curve [20]. The solved optimization problem using linear programs that maximize the detection rates while satisfying the constraints of false positive rates [19].

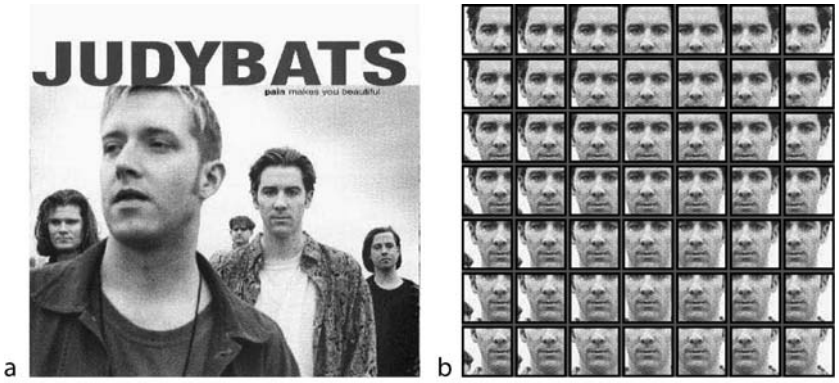
Although the original four types of Haar-like features are sufficient to encode upright frontal face images, other types of features are essential to represent more complex patterns (e.g., faces in different pose) [10, 11, 12, 15]. Most systems take a divide-and-conquer strategy and a face detector is constructed for a fixed pose, thereby covering a wide range of angles (e.g., yaw and pitch angles). A test image is either sent to all detectors for evaluation, or to a decision module with a coarse pose estimator for selecting the appropriate trees for further processing. The ensuing problems are how the types of features are constructed, and how the most important ones from a large feature space are selected. More generalized Haar-like features are defined in [10, 11] in which the rectangular image regions are not necessarily adjacent, and furthermore the number of such rectangular blocks is randomly varied [10]. Several greedy algorithms have been proposed to select features efficiently by exploiting the statistics of features before training boosted cascade classifiers [17].

There are also other fast face detection methods that demonstrate promising results, including the component-based face detector using Naive Bayes classifiers [4], the face detectors using support vector machines [7, 21, 22], the Anti-face method [23] which consists of a series of detectors trained with positive images only, and the energy-based method [24] that simultaneously detects faces and estimates their pose in real time.

Quantifying Performance

There are numerous metrics to gauge the performance of face detection systems, ranging from detection frame rate, false positive/negative rate, number of classifier, number of feature, number of training image, training time, accuracy, and memory requirements. In addition, the reported performance also depends on the definition of a "correct" detection result [2, 5]. Figure 2 shows the effects of detection results versus different criteria, and more discussions can be found in [2, 5].

The most commonly adopted method is to plot the ▶ ROC curve using the de facto standard MIT + CMU data set [2] which contains frontal face images. Another data set from CMU contains images with faces that vary in pose from frontal to side view [4]. Note that



Face Detection. **Figure 2** Detection results depend heavily on the adopted criteria. Suppose all the sub-images in (**b**) are

although the face detection methods nowadays have impressive real-time performance, there is still much room for improvement in terms of accuracy. The detected faces returned by state-of-the-art algorithms are often a few pixels (around 5) off the "accurate" locations, which is significant as face images are usually standardized to 21 × 21 pixels. While such results are the trade-offs between speed, robustness, and accuracy, they inevitably degrade the performance of any biometric applications using the contents of detected faces. Several post-processing algorithms have been proposed to better locate faces and extract facial features (when the image resolution of the detected faces is sufficiently high) [25].

Applications

As face detection is the first step of any face processing system, it finds numerous applications in face recognition, face tracking, facial expression recognition, facial feature extraction, gender classification, clustering, attentive user interfaces, digital cosmetics, biometric systems, to name a few. In addition, most of the face detection algorithms can be extended to recognize other objects such as cars, humans, pedestrians, and signs, etc. [5].

Summary

Recent advances in face detection have created a lot of exciting and reasonably robust applications. As most of the developed algorithms can also be applied to other problem domains, it has broader impact than detecting faces in images alone. Future research will focus on improvement of detection precision (in terms of location), online training of such detectors, and novel applications.

Related Entries

- ▶ Biometric Algorithm
- ▶ Ensemble Learning
- ▶ Face Tracking
- ▶ Face Recognition, Overview
- ▶ Facial Expression Recognition
- ▶ Machine-Learning
- ▶ Supervised Learning Surveillance

References

1. Sung, K.K., Poggio, T.: Example-based learning for view-based human face detection. IEEE Trans. Pattern Anal. Mach. Intell. **20**(1), 39–51 (1998)
2. Rowley, H., Baluja, S., Kanade, T.: Neural network-based face detection. IEEE Trans. Pattern Anal. Mach. Intell. **20**(1), 23–28 (1998)
3. Viola, P., Jones, M.: Robust real-time face detection. Int. J. Comput. Vision **57**(2), 137–154 (2004)
4. Schneiderman, H., Kanade, T.: Object detection using the statistics of parts. Int. J. Comput. Vision **56**(3), 151–177 (2004)
5. Yang, M.H., Kriegman, D., Ahuja, N.: Detecting faces in images: A survey. IEEE Trans. Pattern Anal. Mach. Intell. **24**(1), 34–58 (2002)
6. Mohan, A., Papageorgiou, C., Poggio, T.: Example-based object detection in images by components. IEEE Trans. Pattern Anal. Mach. Intell. **23**(4), 349–361 (2001)
7. Heisele, B., Serre, T., Poggio, T.: A component-based framework for face detection and identification. Int. J. Comput. Vision **74**(2), 167–181 (2007)
8. Fleuret, F., Geman, D.: Coarse-to-fine face detection. Int. J. Comput. Vision **41**(12), 85–107 (2001)
9. Papageorgiou, C., Poggio, T.: A trainable system for object recognition. Int. J. Comput. Vision **38**(1), 15–33 (2000)
10. Dollar, P., Tu, Z., Tao, H., Belongie, S.: Feature mining for image classification. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2007)
11. Li, S., Zhang, Z.: Floatboost learning and statistical face detection. IEEE Trans. Pattern Anal. Mach. Intell. **28**(9), 1112–1123 (2004)
12. Huang, C., Ai, H., Li, Y., Lao, S.: High-performance rotation invariant multiview face detection. IEEE Trans. Pattern Anal. Mach. Intell. **29**(4), 671–686 (2007)
13. Freund, Y., Schapire, R.: A decision-theoretic generalization of on-line learning and application to boosting. J. Comput. Syst. Sci. **55**(1), 119–139 (1997)
14. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). Ann. Stat. **28**(2), 337–407 (2000)
15. Jones, M., Viola, P.: Fast multi-view face detection. Technical Report TR2003-96, Mitsubishi Electrical Research Laboratories (2003)
16. Viola, P., Jones, M.: Fast and robust classification using asymmetric Adaboost and a detector cascade. In: Advances in Neural Information Processing Systems, pp. 1311–1318 (2002)
17. Wu, J., Brubaker, S.C., Mullin, M., Rehg, J.: Fast asymmetric learning for cascade face detection. IEEE Trans. Pattern Anal. Mach. Intell. **30**(3), 369–382 (2008)
18. Pham, M.T., Cham, T.J.: Online learning asymmetric boosted classifiers for object detection. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2007)
19. Brubaker, S.C., Wu, J., Sun, J., Mullin, M., Rehg, J.: On the design of cascades of boosted ensembles for face detection. Int. J. Comput. Vision **77**(1–3), 65–86 (2008)
20. Provost, F., Fawcett, T.: Robust classification for imprecise environments. Mach. Learn. **42**(3), 203–231 (2001)

21. Oren, M., Papageorgiou, C., Sinha, P., Osuna, E., Poggio, T.: Pedestrian detection using wavelet templates. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 193–199 (1997)
22. Romdhani, S., Torr, P., Schölkopf, B., Blake, A.: Computationally efficient face detection. In: Proceedings of the Eighth IEEE International Conference on Computer Vision, vol. 2, pp. 695–700 (2001)
23. Keren, D., Osadchy, M., Gotsman, C.: Antifaces: A novel fast method for image detection. IEEE Trans. Pattern Anal. Mach. Intell. **23**(7), 747–761 (2001)
24. Osadchy, M., LeCun, Y., Miller, M.: Synergistic face detection and pose estimation with energy-based models. J. Mach. Learn. Res. 1197–1214 (2007)
25. Ding, L., Martinez, A.: Precise detailed detection of faces and facial features. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2008)

Face Device

Massimo Tistarelli
Computer Vision Laboratory, University of Sassari, Piazza Duomo, 6 Alghero, Italy

Synonyms

Face acquisition; Face camera; Video camera; Visual sensor

Definition

A face device is a system to acquire a set of digital data samples representing a human face. As the human face is a complex 3D object, the data can be in several forms: a 2D image where the gray levels of the ▶ pixels represent the projected reflectance of the face surface under visible illumination; a 2D image where the gray levels of the pixels represent the projected reflectance of the face surface illuminated with an active source; a 2D thermal image representing the heat emitted by the face surface; 3D samples of the surface structure.

Face devices can be distinguished on the basis of the data dimension if it is active or passive. Face devices can be passive, i.e., based on the passive reflectance of ambient light by the body, or active, i.e., associated with an energy emitter and a sensor to capture the energy reflected by the face. The data captured can be either in 2D or 3D form.

A face device can be based on different technologies, depending upon the data to be captured and the signal to be obtained. The most applied face devices include a video ▶ camera to capture 2D images of the face and a digitizer to sample and quantize the analog signal generated by the camera. Different face devices deliver different signals to be digitized into 2D or 3D data. The data captured can be stored under different file formats for subsequent processing.

Introduction

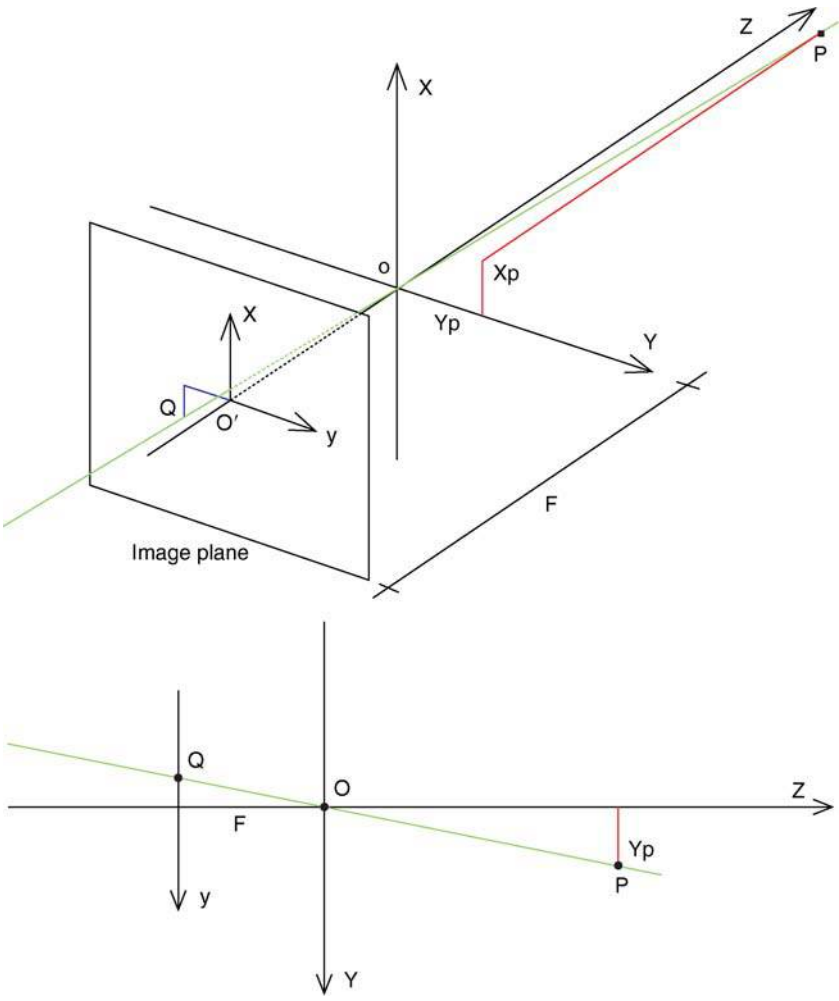
Current face biometric systems are based on the acquisition and processing of image data, representing a human face. A face acquisition device is typically a video camera capable of acquiring single images or video streams of data, representing a face. As the face is a 3D object, the acquired data can represent either the 2D projection of the face reflectance on the image plane or a set of 3D samples of the face structure, possibly with the associated reflectance. In the former case, a conventional video camera can be used to acquire images of face. In the latter case, a more complex 3D acquisition device must be applied.

2D Face Devices

A conventional camera acquires the image data as a reflectance of the imaged scene. The face points are recorded as the geometrical projection of the 3D points on the face surface onto the 2D image plane (Fig. 1).

Several video cameras exist that are capable of capturing either single images or video streams from the viewed scene. The most critical parts of the camera are the acquisition sensor and the lenses.

Charge-coupled device (CCD) and complementary metal oxide semiconductor (CMOS) image sensors are two different technologies for capturing images digitally; current commercial camera adopt either of these. Both types of sensors convert light into electric charge and process it into electronic signals. In a ▶ CCD sensor, charge of every pixel is transferred through a very limited number of output nodes



Face Device. **Figure 1** Geometry of the pin hole camera model. (Top) 3D sketch of the projection of point *P* in space on the image pixel *Q*. (Bottom) 2D projection of the *Y–Z* plane.

(often just one) to be converted to voltage, buffered, and sent off-chip as an analog signal. All the pixel can be devoted to capture light, and the output's uniformity (a key factor in image quality) is high. In a ► CMOS sensor, each pixel has its own charge-to-voltage conversion, and the sensor often includes amplifiers, noise-correction, and digitization circuits, so that the chip outputs digital bits. These other functions increase the design complexity and reduce the area available for light capture. With each pixel doing its own conversion, the uniformity is low. But the chip can be built to require less off-chip circuitry for basic operation.

The CMOS pixel solves the speed and scalability issues of the CCD sensor. They consume far less power than a CCD, have less image lag, and can be fabricated on much cheaper and more available manufacturing lines. Unlike CCDs, CMOS sensors can combine both the image sensor and the image processing functions within the same integrated circuit. CMOS imagers still suffer from higher ► fixed-pattern noise than CCDs, but active pixel sensors are catching up with respect to noise, dynamic range, and responsivity. CMOS sensors have become the technology of choice for many consumer applications, most significantly, the burgeoning cell phone camera market [1].

The technology of the sensor and the capturing device determines several properties of the captured signal such as the following:

1. *Image resolution*. This is related to both the active elements on the imager sensor and the sampling

device used to digitize the signal. Even though solid state sensors are used in digital cameras, they produce an analog video signal. As a consequence, the captured image resolution strongly depends on the sampling frequency of the digitization device. Other factors affecting the image resolution are the file standard format adopted for the image storage and the image processing application required to postprocess the face images.

2. *Responsivity*. The amount of signal the sensor delivers per unit of input optical energy. CMOS imagers are marginally superior to CCDs, in general, because gain elements are easier to be place on a CMOS image sensor. This affects the illumination level required to capture a face image with a sufficient contrast level.
3. *Dynamic range*. The ratio of a pixel's saturation level to its signal threshold. CCD sensors are much better than CMOS in this regard. Some CMOS sensors deliver 8 bit per pixel intensities, corresponding to 128 real level variations. As a consequence, the information content in the image features is half than what is expected. A higher dynamic range implies a higher image contrast even at low illumination levels and the possibility to grab finer details. A gray level quantization of 8 bit per pixel is generally sufficient for capturing good quality face images. The sensor dynamic range can be crucial when acquiring color images. In this case, the color quantization may influence the information content in the face image itself, especially if a low bit rate (with less than 8 bit per color channel) is used for color coding.
4. *Sensitivity to noise* (signal to noise ratio – SNR). The three primary broad components of noise in a CCD imaging system are photon noise (results from the inherent statistical variation in the arrival rate of photons incident on the CCD), dark noise (arises from statistical variation in the number of electrons thermally generated within the silicon structure of the CCD), and read noise (a combination of system noise components inherent to the process of converting CCD charge carriers into a voltage signal for quantification, and the subsequent processing including the analog-to-digital (A/D) conversion). A further useful classification distinguishes noise sources on the basis of whether they are temporal or spatial. CCDs still enjoy significant noise advantages over CMOS imagers because of quieter sensor substrates (less on-chip circuitry), inherent tolerance to bus capacitance variations, and common output amplifiers with transistor geometries that can be easily adapted for minimal noise.
5. *Uniformity*. The consistency of response for different pixels under identical illumination conditions. Spatial wafer processing variations, particulate defects, and amplifier variations create nonuniformities in light responses. It is important to make a distinction between uniformity under illumination and uniformity at or near dark. CMOS imagers were traditionally much worse than CCDs under both regimes. New on-chip amplifiers have made the illuminated uniformity of some CMOS imagers closer to that of CCDs, sustainable as geometries shrink. This is a significant issue in high-speed applications, where limited signal levels mean that dark nonuniformities contribute significantly to overall image degradation.
6. *Shuttering*. The ability to start and stop exposure arbitrarily. It is a standard feature of virtually all consumer and most industrial CCDs, especially interline transfer devices, and it is particularly important in machine vision applications. CCDs can deliver superior electronic shuttering, with little fill-factor compromise, even in small-pixel image sensors. Implementing uniform electronic shuttering in CMOS imagers requires a number of transistors in each pixel. In line-scan CMOS imagers, electronic shuttering does not compromise fill factor, because shutter transistors can be placed adjacent to the active area of each pixel. In area-scan (matrix) imagers, uniform electronic shuttering comes at the expense of fill factor, because the opaque shutter transistors must be placed in what would otherwise be an optically sensitive area of each pixel. A uniform synchronous shutter, sometimes called a nonrolling shutter, exposes all pixels of the array at the same time. Object motion stops with no distortion, but this approach reduces the pixel area because it requires extra transistors in each pixel. Users must choose between low fill factor and small pixels on a small, less-expensive image sensor, or large pixels with much higher fill factor on a larger, more costly image sensor.
7. *Sampling speed*. This is an area in which CMOS arguably delivers better performances over CCDs,

because all camera functions can be placed on the image sensor. With one die, signal and power trace distances can be shorter, with less inductance, capacitance, and propagation delays. To date, CMOS imagers have established only modest advantages in this regard, largely because of early focus on consumer applications that do not demand notably high speeds compared with the CCD's industrial, scientific, and medical applications. Both the sampling and shuttering speed are important when capturing video streams of faces. In this case, it is important to ensure the image stability and minimize the motion smear induced by either the motion of the camera or the face. This requires to tune the camera sampling frequency to the motion speed induced in the image sequence. If the face is very close to the camera, small motions can induce large and fast displacements on the image, thus producing motion smear. At a larger distance (above 50 cm), a standard sampling frequency of 50 or 60 Hz is generally sufficient. In many low-cost devices, the sampling frequency depends on the time required to transmit the signal from the device to the frame buffer. Therefore, only low resolution images can be captured at high sampling frequencies. On the other hand, if a high, nonstandard sampling frequency is required to capture stable images with fast motions, the reduced exposure time requires a higher sensitivity of the sensor to preserve a high SNR.

8. *Windowing*. One unique capability of CMOS technology is the ability to read out a portion of the image sensor. This allows elevated frame or line rates for small regions of interest. This is an enabling capability for CMOS imagers in some applications, such as high-temporal-precision face tracking in the subregion of an image. CCDs generally have limited abilities in windowing.
9. *Antiblooming*. The ability to gracefully drain localized overexposure without compromising with the rest of the image in the sensor. CMOS generally has natural blooming immunity. CCDs, on the other hand, require specific engineering to achieve this capability; many CCDs that have been developed for consumer applications do, but those developed for scientific applications generally do not.
10. *Biasing and noise*. CMOS imagers have a clear edge in this regard. They generally operate with a single bias voltage and clock level. Nonstandard biases are generated on-chip with charge pump circuitry isolated from the user unless there is some noise leakage. CCDs typically require a few higher-voltage biases, but clocking has been simplified in modern devices that operate with low-voltage clocks.

The camera optics determines the general image deformation, the depth of the field, and the amount of blurring in the image. The lenses must be chosen carefully according to the acquisition scenario. The ▶ focal length must be set to provide a sufficient ▶ depth of field (DOF) to always keep the subject's face in focus. If the range of distances is very large, a motorized lens can be used to dynamically keep the face in focus. Otherwise, a shorter focal length lens, with a larger depth of field, can be used at the expenses of an increase in the image distortion.

A 2D camera can be modeled with several parameters [2], including the following:

1. The (X, Y, Z) position of the center of the camera lens
2. The focal length
3. The orientation of the sensor's plane
4. The aperture or ▶ field of view (Xf, Yf)
5. The physical x and y dimensions of each pixel on the sensor
6. The normal to the focal plane
7. The lenses properties

Many parameters can be neglected in the *pin hole* camera model. This is a simplified model where the physical parameters are reduced to five virtual parameters, namely the following:

1. The focal length F
2. The pixel width and height δx, δy
3. The x and y coordinates of the optical center (x_c, y_c)

Assuming the pin hole camera model, the (x, y) projection on the image plane of a 3D point (X, Y, Z) can be represented as (refer to Fig. 1)

$$\begin{aligned} x &= F_x \frac{X}{Z}, \\ y &= F_y \frac{Y}{Z}, \end{aligned} \tag{1}$$

where F_x and F_y represent the two values of the focal length, which take into account the image aspect ratio.

The pin hole model cannot take into account several effects of the misalignment of the sensor with the lenses, not the lens aberration or the image deformation due to the focal length. However, when high accuracy is required or when low-end cameras are used, additional effects have to be taken into account.

The failure of the optical system to bring all light rays received from a point object to a single image point or to a prescribed geometric position should then be taken into account. These deviations are called aberrations. Many types of aberrations exist (e.g., astigmatism, chromatic aberrations, spherical aberrations, coma aberrations, curvature of field aberration, and distortion aberration). It is outside the scope of this work to discuss them all. The interested reader is referred to the work of Willson [3] and to the photogrammetry literature [4].

Many of these effects are negligible under normal acquisition circumstances. Radial distortion, however, can have a noticeable effect for shorter focal lengths. Radial distortion is a linear displacement of image points radially to or from the center of the image, caused by the fact that objects at different angular distance from the lens axis undergo different magnifications. It is possible to cancel most of this effect by ► Face Warping the image.

Active 2D Face Devices

Within the general class of 2D face devices, active devices rely on the possibility to use an active source of energy to radiate the subject's face. Among them, the most commonly used are the near infrared cameras. These cameras have normal optics but the sensor (either CMOS or CCD) is sensitive to a wavelength spectrum between 0.7 and 1.1 μm. To perform image acquisition, the subject's face to be captured must be illuminated by an infrared illuminator. Given the sensitivity response curve of the near infrared sensor, the pixel intensities are almost exclusively due to the reflection of the infrared light on the face skin. An example is presented in Fig. 2. As a consequence, a remarkable advantage of this face acquisition device is the relative insensitivity to changes in environmental illumination.

Face Device. **Figure 2** Sample image acquired with a near-infrared camera.

3D Face Devices

Another category of face device are those aimed at acquiring the 3D shape information of the face. There are several technologies applied to produce 3D cameras for face acquisition. They can be broadly grouped in the following categories:

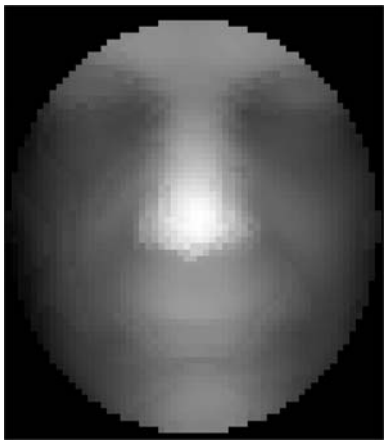
1. *Stereo triangulation cameras.* A pair of stereo cameras is used for determining the depth to points on the face, for example, from the center point of the line between their focal points. To solve the depth measurement problem using stereo cameras, it is necessary to first find corresponding points in the two images. Solving the correspondence problem is one of the main problem when using this type of technique. As a consequence, range imaging based on stereo triangulation can usually produce reliable depth estimates only for a subset of all points visible in both cameras. The advantage of this technique is that the measurement is more or less passive; it does not require special arrangements in terms of scene illumination.
2. *Light stripe triangulation.* Illuminating the face with a light stripe creates a reflected line as seen from the light source. From any point out of the plane of the stripe, the line will typically appear as a curve, the exact shape of which depends both on the distance between the observer and the light source and on the distance between the light source

and the reflected points. By observing the reflected sheet of light using a camera (often a high resolution camera) and knowing the positions and orientations of both camera and light source, it is possible to determine the distance between the reflected points and the light source or camera. By moving either the light source (and normally also the camera) or the scene in front of the camera, a sequence of depth profiles of the scene can be generated. These can be represented as a 2D range image. The most common cameras are based on the projection of an invisible and unharmful laser light stripe. The light stripes projected along the face surface are captured by a conventional camera. The distortion in the light stripes induced by the face shape is computed to infer the 3D structure of the surface.

3. *Time-of-flight laser scanner.* The time-of-flight 3D laser scanner is an active scanner that uses laser light to probe the subject. At the heart of this type of scanner is a time-of-flight laser range finder. The laser range finder finds the distance of a surface by timing the round-trip time of a pulse of light. A laser is used to emit a pulse of light and the amount of time before the reflected light is seen by a detector. Since the speed of light c is known, the round-trip time determines the travel distance of the light, which is twice the distance between the scanner and the surface.

Inspite of the camera and sensor technology, the produced image is either a depth map, a collection of 3D points in space, or a set of 3D features representing the 3D structure of the acquired face. A sample depth map of a face is shown in Fig. 3. The most frequently used representations for the acquired 3D data can be listed as follows:

1. *Point cloud.* A large number of 3D points that are sampled from the surface of the face are stored.
2. *3D mesh.* Triangulation is used to produce a mesh from the point cloud. This is a more compact representation. Range images – One or more 2D range images can be stored, especially if the range data are taken from a single perspective.
3. *Feature sets.* There are different features that one can derive and store for each face. Typical features are landmark locations (nose tip, eyes, corners of the mouth, etc.), surface normals, curvatures, profile features, shape indices, depth and/or colour histograms, edges, and subspace projection coefficients (PCA and LDA are frequently used).



Face Device. **Figure 3** Sample depth face image. The gray levels are inversely proportional to the distance of the face surface from the camera [5].

The point cloud representation is the most primitive 3D information provided by a 3D camera. The 3D-RMA is an example of a database of 3D face models represented by clouds of points [6]. For long time, it has been the only publicly available database, even if its quality is rather low. Meshes are obtained by triangulation. These are more structured and easier to deal with. Data in the form of meshes are more available today, but in most cases the mesh databases are proprietary. Usually, more than one representation is used in a single algorithm. Texture information, if available, is generally stored for each 3D point or triangle. A sample 3D face image with the associated reflectance map is shown in Fig. 4.

Summary

A face device is a system to acquire a set of digital data samples representing a human face. As the human face is a complex 3D object, the data can be in several forms, from a 2D image to a complex 3D representation. The principal component of a face device is a digital camera, which acquires images either for a direct 2D representation or to build a 3D representation of the face shape. Different cameras offer variable performances, in terms of quality of the signal,

Face Device. Figure 4 Sample 3D face image and projected 2D intensity values from the face recognition grand challenge (FRGC) [5] database.

sensitivity to different light spectral components, and capturing speed. The proper imaging device must be carefully chosen for the application scenario. The ambient illumination level, the required level of detail, the effects of noise, and the motion speed of the objects in the scene must all be carefully considered.

Related Entries

- ▶ Acquisition
- ▶ Authentication
- ▶ Enrollment
- ▶ Identification
- ▶ Verification

References

1. Litwiller, D.: CCD vs. CMOS: facts and fiction. Photonics Spectra, pp. 151–154 (2001)
2. Blais, F.: Review of 20 years of range sensor development. J. Electron. Imaging **13**(1), 231–240 (2004)
3. Willson, R., Shafer, S.: What is the center of the image? J. Opt. Soc. Am. A **11**(11), 2946–2955 (1994)
4. Slama, C.: Manual of Photogrammetry/ American Society of Photogrammetry, Falls Church, VA, USA, 4th edn. (1980)
5. Phillips, J.J., Flynn, P., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Jaesik, M., Worek, W.: Overview of the face recognition grand challenge. In Proceedings *CVPR05,* pp. 947–954 (2005)
6. Beumier, C., Acheroy, M.: Automatic 3D face authentication. Image Vision Comput. **18**(4), 315–321 (2000)

Face Identification

- ▶ Face Recognition, Thermal
- ▶ Forensic Evidence of Face

Face Image Data Interchange Formats

- ▶ Face Image Data Interchange Formats, Standardization

Face Image Data Interchange Formats, Standardization

Patrick Grother, Elham Tabassi
National Institute of Standards and Technology, USA

Synonym

Face image data interchange formats

Definition

Openly documented data structures for universally interpretable interchange of facial imagery.

Biometric data interchange standards are needed to allow the recipient of a data record to successfully process data from an arbitrary producer. This defines biometric interoperability and the connotation of the phrase "successfully process" is that the sample, in this case, a facial image record, can be accurately identified or verified. This can be achieved only if the data record is both syntactically and semantically conformant to a documentary standard.

Introduction

Facial image standards are perhaps the oldest documented biometric data standards. Predating even the fingerprint, the facial image has been mandated for identity documents since at least the World War I when several European governments saw the need for a facial photograph to serve as the core element in the cross-border identity verification application. Of course the data record was simply an analog paper printed photograph - the advent of fully automatic face recognition algorithms and the need for digital images was at least 70 years distant [1, 2]. However the intention remains the same: to support (human or machine) verification of an individual via a high quality standardized image.

Roles

The use of face imagery for recognition is ubiquitous in applications where a human does the recognition. This rests on three factors: The ability of humans to recognize faces; the almost universal availability of the face In some cultures the face is covered or painted, and in such cases modalities such as iris or hand geometry are dominant.; and the availability of cameras and printers. The result is that face images, printed on passports, drivers' licenses, credit cards, and other tokens, have been the primary biometric element for human verification for many years.

Nowadays with the advent and maturation of technologies for automated face recognition, the use of the face for verification [3] is but one component of a larger marketplace in which commercial systems have been both piloted and fully deployed for identification applications such as watch-list surveillance [4] and duplicate detection (e.g., for drivers licenses, or visas). In addition the law enforcement community has for years taken mugshot images and, while these are often only used for human identification, they are being used operationally [5].

The common theme among all is that recognition accuracy is critically sensitive function of the quality of the image, where quality here refers to the photometric and geometric properties of the image. The former include contrast, exposure, and uniformity of lighting; the latter refers to the size of the image and the angular orientation of the face to the viewing direction. The effect of non-idealities in these areas has been quantified extensively and there is an enormous literature documenting research in how to improve the robustness and invariance of the algorithms to variations in these quantities. In parallel, there has been a concerted effort by groups of vendors, users, governmental organizations, and academics to develop standards that establish a baseline for the acquisition and quality of the captured images.

It is no coincidence that the largest marketplace for face recognition technologies today is in those applications where the quality is most highly controlled, namely passports and visas, where the photographers and the subjects who pay them, are positively motivated to provide good conformant images.

In a more general sense, formal face images standards also serve to do what many other data format standards do: they define a parseable record that allows syntactic interoperability. This creates a foundation for a marketplace of off-the-shelf products, and is a necessary condition to achieve supplier independence, and to avoid vendor lock-in. It is perhaps surprising that in a world where many raster image formats are open and standardized [6–8] it remains common for images to be retained in a fully proprietary (i.e., unpublished) format. Such practice may be acceptable *within* an application, but is a serious impediment once cross-organizational interchange of data is required. This is the essence of interoperability which allows modular integration of products without compromising architectural scope, and it facilitates the upgrade process and thereby mitigates against obsolescence.

The business implications of these benefits are many. A good standard, well implemented, may create entirely new markets (e.g., e-Passports include face image records). On the other hand, robust standards tend to lead to competition and reduced profit margins. This process, commoditization, is an inhibitory factor for many technology companies that balance the

promise of new or expanded marketplaces against reduced barriers to entry for competitors. The decision is determined by the amount of intellectual property that a standard allows suppliers to hide behind its implementation. From the user perspective, standards may serve to enhance competition and performance. For example, face image standards (primarily ISO/IEC 19794-5 [9]), which are currently being mandated in a number of large government and international programs, specify image formats without requiring particular equipment or matching algorithms.

This is the motivation for formal published consensus standards.

Standards do not in and of themselves assure interoperability. Specifically, when a standard is not fully prescriptive, or it allows for optional content, then two implementations that are both exactly conformant to the standard may still not interoperate. This situation may be averted by applying further constraints on the application of the standard. This is done by means of "application profile" standards which formally call out the needed base standards and refine their optional content and interpretation.

History of Face Standardization

The current face standards descend from standardization efforts that began in the mid 1990s. These were driven in large part by the needs of the United States' Federal Bureau of Investigation who sought to establish uniform standards for State and local law enforcement authorities submitting images to them.

Referring to Table 1, the first standard, approved in April 1997, established the syntax of a record denoted "Type 10." The image data it required was either in raw grayscale format or, if compressed, in the then draft JPEG/JFIF standard [6]. Concurrently NIST established procedures for the geometric and photometric properties of images and published its recommendations in September 1997. These were extended and modified, and incorporated, in 2000, into both the American Association of Motor Vehicle Administrators standard for drivers licenses, and the revision of the FBI's original biometric data specifications.

These standards formed the basis for the subsequent development of the national INCITS 385:2004 standard in 2004, which in turn begat the full ISO/IEC 19794-5 International Standard in 2005. (At the time of writing the standard is under amendment to regulate the acquisition process, and to establish a container for three dimensional data.) A substantially revised standard which would include these changes (and others) is likely to be completed late in the decade.

The ISO/IEC 19794-5 Face Image Standard

The ISO/IEC 19794-5:2005 standard is the fifth part of a multipart biometric data interchange format standard. The standard is organized by modality, and other parts cover fingerprint images, irises, and hand geometry among many others. The Part 5 standard is the most widely implemented, most actively developed, and most modern face standard. Its content drove the

Face Image Data Interchange Formats, Standardization. Table 1 The evolution of contemporary face image standards

Date	Title of Standard
04/1997	Addendum To ANSI/NIST-CSL 1-1993 (adding Mugshots, scars, marks and tattoos)
09/1997	NIST Best Practice Recommendation for the Capture of Mugshots
06/2000	AAMVA National Standard for the Driver License/Identification Card
09/2000	ANSI/NIST-ITL 1-2000 - Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo Information - Type 10
05/2004	INCITS 385:2004 - Face Recognition Format for Data Interchange
06/2005	ISO/IEC 19794-5:2005 - Face Image Data
04/2007	ANSI/NIST-ITL 1-2007 - Data Format for the Interchange of Fingerprint, Facial, & Other Biometric Information - Type 10
06/2007	ISO/IEC 19794-5/Amd 1 - Conditions for Taking Photographs for Face Image Data
2009 (Est)	ISO/IEC 19794-5/Amd 2 - Three Dimensional Face Image Data Interchange Format

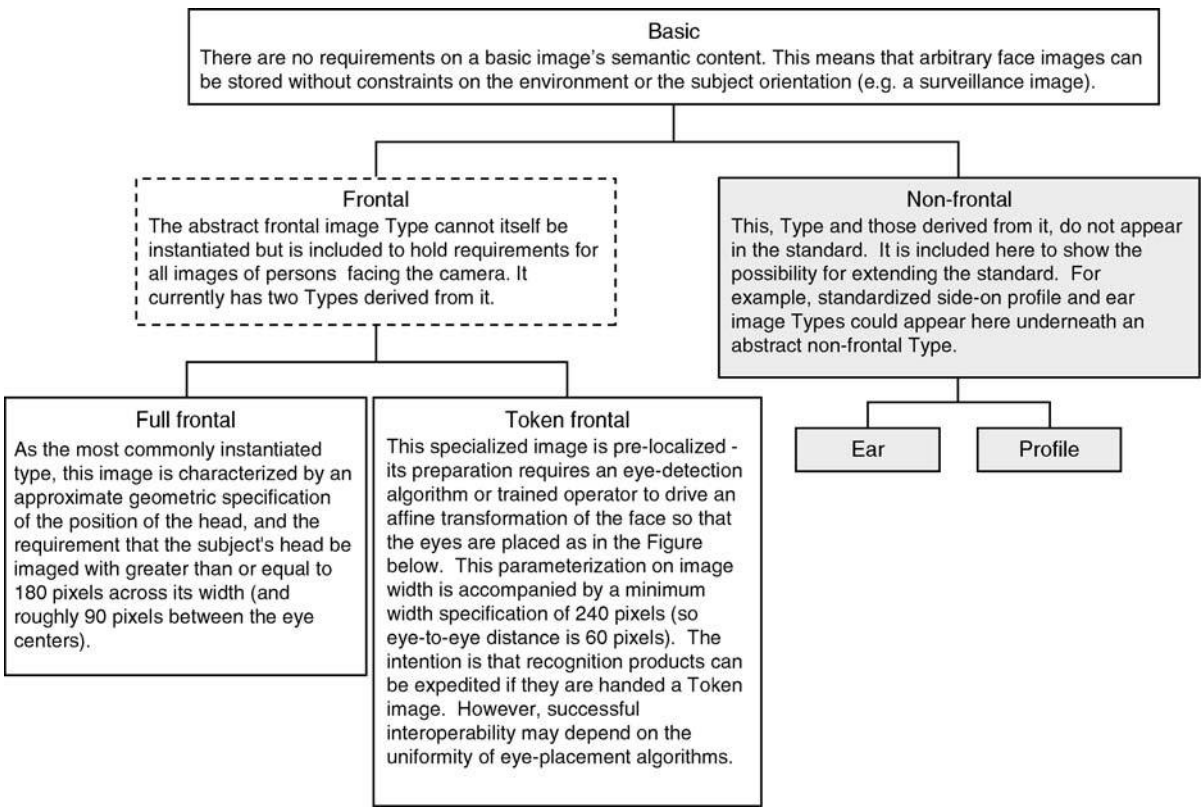
revision of the Type 10 record of the ANSI/NIST ITL 1-2007 described in section. While the ISO standard is under revision, with publication due late in the decade, the existing 2005 standard has been called out for some major identity management applications. The foremost of these is the e-Passport, which the International Civil Aviation Organization formalized in its ICAO 9303 standard. This points to ISO/IEC 19794-5 as the mandatory globally interoperable data element for ISO/IEC 14443 contactless chip passports.

The face standard defines a binary record structure for the storage of one or more face images. It establishes requirements on the syntax and semantic content of the structure. These requirements are stated in terms of the following four categories.

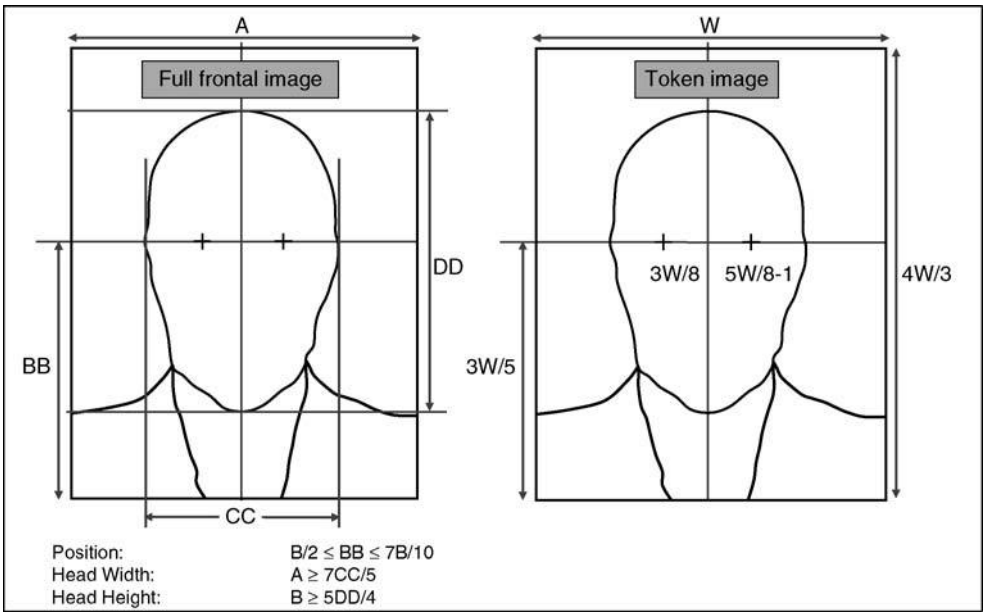
- *Format*: These requirements detail the syntactic arrangement of the data elements in the record.
- *Scene*: These requirements regulate variables such a pose, expression, shadows on the face, the wearing of eye glasses.
- *Photographic*: These requirements concern correct exposure of the subject, distortion, focus, and depth of field.
- *Digital*: The requirements include specifications for dynamic range, color space, pixel aspect ratio, and video interlacing.

The standard imposes these requirements incrementally: Fig. 1 shows that the useful frontal image types inherit from parent types and add requirements. This object oriented design allows for future specialized types to be added, including 3D frontal types. In addition the standard establishes two geometric position specifications for the face. These are shown in Fig. 2. The tighter specification, for known as the token Frontal, requires detection of the eye coordinates and of fine transformation of the image.

The record includes fields for expression, eye-color, hair color, and gender. It optionally allows the inclusion of ISO/IEC 14496-2 MPEG 4 feature points. The standard includes various quality related requirements. For example the pose angle is required to be $\pm$ 5*deg*, and there must be at least 7 bits of greylevel information on the face. Conformance to these requirements will elevate face recognition performance. Once an image is acquired a test of its conformance to the standard's specifications requires some non-trivial



Face Image Data Interchange Formats, Standardization. **Figure 1** Inherited types of the ISO/IEC 19794-5 face image standard.



Face Image Data Interchange Formats, Standardization. **Figure 2** Geometries of the ISO/IEC 19794-5 frontal face images.

image analyses. A number of software products have been developed to "box-check" ISO conformance and to prepare the standardized record.

The ANSI/NIST ITL 1-2007 Type 10 Record

Since its initial development in the early 1990s, the so-called ANSI-NIST standard has been very widely implemented and used within and between the law enforcement communities of the United States and the many other countries. Its primary use is the transmission of fingerprint data from the State and Local authorities to central automated fingerprint identification systems, primarily those operated by the Federal Bureau of Investigation. The ANSI/NIST standard includes defined *Types* for the major biometric modalities. The standard is multimodal in that it allows a user to define a transaction that would require, for example, fingerprint data as Type 14, a facial mugshot as Type 10, and the mandatory header and metadata records Type 1 and 2. These are linked with a common numeric identifier.

Of concern here, since its development in 1997, is the Type 10 record. It supports storage not just of face images, but also those of scars, marks, and tattoos, with the particular type of content being recorded in the "image type" field of the header.

Unlike the ISO standard's fixed binary structure, the Type 10 has a tag-value structure in which a three letter code begins a field. The mandatory fields are: Record length, image designation code (identifier linking, say, Type 14 finger + Type 10 face records), image type (face or otherwise), the source agency (e.g., local police department), the capture date, the width, height and scanning resolution, the compression algorithm, color space, and the subject acquisition profile. This latter field encodes, essentially, the conformance of the image to particular capture specifications. These are either established elsewhere [9–12] or introduced in the standard.

The optional fields are: pose category (frontal, profile, other), actual pose angles, whether the subject was wearing headwear or eyewear, the camera type, a quality value and it's source, the eye and hair color, facial expression, eye and nostril locations and MPEG 4 feature points, and whether the capture was attended or automatic. The last field contains the image data itself, which is either an uncompressed raw greyscale or color image, or a JPEG, JPEG 2000 or PNG encoded image.

Amendment 1 to ISO/IEC 19794-5:2005

The 2007 amendment is an informative Annex to the base 2005 face standard. It is written to provide expert

guidance for the photography of faces particularly by owners and operators of studios, photo stores or other organizations producing or requiring either printed photographs or digital images that would conform. It is intended to assist in the production of images that are conformant to the frontal type requirements of the base standard.

The standard regulates the subject, lighting, and camera placement for three kinds of face acquisition environments listed here in the order of increasing space constraints and non-ideality: a photo studio (e.g., for a passport), a registration desk (e.g., for a driving license), and a photo-booth. For each of these the standard addresses camera-subject positioning (in terms of distance, height, focus, and depth of field), exposure (in terms of F-stops and shutter speed), and illumination (in terms of number, type and placement of lights). The document also provides guidance on printing and scanning of paper photographs.

Amendment 2 to ISO/IEC 19794-5:2005

A second amendment is currently under preparation. This is aimed at standardizing a container and specifications for images that include three dimensional shape information of the human head. An initial effort within the United States, INCITS 385:2004 Amendment 1, allowed a 2D face image to be accompanied by a z-axis range map (e.g., from a structured light sensor). This shape information was recorded as the intensity values in a greyscale PNG image. The ISO standardization process has recently sought to allow more complete 3D information including the ability to encode concavities and folded structures (e.g., hook nose).

The standards are also likely to allow the storage of 3D information computed from 2D information such as morphable models [13] and active appearance models [14].

Resolution Requirements

The image sizes mentioned in ISO/IEC 19794-5:2005 are very much less than those attainable with contemporary consumer grade digital cameras. The reasons for this are two. First, the face recognition algorithms of the early part of the decade were designed to operate with an interocular eye distance of between 40 and maybe 120 pixels. Second, the standard aims to be application independent, i.e., to only establish a minimum resolution to support automated face recognition. While more modern implementations are capable of exploiting high resolution imagery, the images may be too large for operational use (e.g., on an e-Passport chip, where size is typically much lesser than 50KB). Nevertheless, the 2007 revision of the ANSI/NIST ITL 1-2007 standard reflected the utility of high resolution imagery by incorporating a laddered scale that culminates in an image with a width such that 1700 or more pixels lie on the faces of 99% of U.S. male subjects. This specification supports forensic analysis of the face. It is termed Level 51 and is the highest level of the Type 10 record's Subject Acquisition Profile stack.

Note that a separate *profile* standard or requirements document could normatively specify minimum or maximum resolutions for a particular application. For example, the PIV specification[11] requires that imaging of a 20cm target at 1.5 metres produces 240 pixels, corresponding to about 90 pixels between the eyes.

Note that no standard currently exists for the certification of face recognition imaging systems. Such a standard might reasonably establish true resolution specifications in terms of line pairs per millimeter and as a full modulation transfer function profile. This would regulate the entire imaging system including the effects, say, of video compression.

Standards Development Organizations

Standards are developed by a multitude of standards development organizations (SDOs) operating in a great variety of technical disciplines. SDO's exist within companies and governments, and underneath trade associations and international body umbrellas. International standards promise to regulate larger marketplaces and the development process involves more diverse and thorough review and so consensus is more difficult to achieve. With stakes often high, development processes are conducted according to definitive sets of rules. These are intended to achieve consensus standards that are legally defensible, implementable, and effective.

The following list gives an overview of the relevant SDOs. Note that the published standards are usually copyrighted documents and available only by purchase.

- *ISO JTC 1 SC 37*: Although face image standardization is underwaywithin a number of SDOs, by far the most work is conducted inthe main international forum, SubCommittee 37 (SC 37) *Biometrics*. This body was established in mid 2002 as the newest of seventeenactive subcommittees under Joint Technical Committee 1 (JTC 1)and its parent the International Organization for Standardization(ISO)ISO maintains a catalog of its standards development effortsat http://www.iso.org/iso/en/CatalogueListPage.CatalogueList. Although its focus is development of standards in support of genericidentity management and security applications, its establishment wassubstantially motivated by a need for improved international bordercrossing mechanisms.

  Within the six working groups of SC 37, the body responsible for facial image standardization is Working Group 3. The group, which develops biometric data interchange format standards, is the largest WG in SC 37 and is developing the standards with the highest profile adoption in the marketplace. Its ISO/IEC 19794-5:2005 face image data standard has been specified by the International Civil Aviation Organization (ICAO) as the mandatory biometric in the electronic Passports now being issued in many developed nations.
- *M1*: M1 is the United States Technical Advisory Group (TAG) to SC 37. It was established in June 2002 and is responsible for formulating U.S. positions in SC 37 where it holds the U.S. vote. Staff from its member organizations represent these positions in SC 37. It is notable because it is also a standards development organization in its own right. Particularly it developed and published the INCITS 385INCITS, which stands for International Committee for Information Technology Standards, is the SDO arm of the Information Technology Industry Council based in Washington DC. face image standard in 2004. This document is substantially similar to the ISO/IEC 19794-5 standard because the early drafts of the former were contributed toward the development of the latter.
- *ANSI/NIST*: The U.S. National Institute of Standards and Technology (NIST) is also a SDO. It developed the ANSI/NIST standards for law enforcement under the canvass process defined by ANSI. (see sec.).

Summary

Data interchange standards have been developed to facilitate universal seamless exchange of facial information. In all cases, these wrap an underlying standardized encoded image (often ISO/IEC 10918 JPEG) with a header that includes subject-specific information and details of the acquisition. The standards support accurate face recognition by constraining the cameras, environment, and the geometric and photometric properties of the image.

Related Entries

- ▶ Face Recognition
- ▶ Interoperability

References

1. Kanade, T.: Picture processing system by computer complex and recognition of human faces. In: Doctoral dissertation, Kyoto University (1973). Available as TIFF images at http://xiotech.ulib.org/cgi-bin/ulib/display?11014.12072
2. Sirovich, L., Kirby, M.: Low dimensional procedure for the characterization of human faces. J. Opt. Soc. Am. A **4**(3), 519–524 (1987)
3. Australian Customs Service: Smartgate. Tech. rep.
4. Face recognition as a search tool "foto-fahndung". Tech. rep.
5. Frank, T.: Face recognition next in terror fight. USA Today (2007)
6. JTC 1, SC29 Coding of audio, picture, multimedia and hypermedia information: ISO/IEC 10918-1 Digital compression and coding of continuous-tone still images: Requirements and guidelines, 1 edn. (1994). URL http://webstore.ansi.org International Standard
7. JTC 1, SC29 Coding of audio, picture, multimedia and hypermedia information: ISO/IEC 15948 Computer graphics and image processing – Portable Network Graphics (PNG): Functional specification, 1 edn. (2004). URL http://webstore.ansi.org International Standard
8. JTC 1, SC29 Coding of audio, picture, multimedia and hypermedia information: ISO/IEC 15444-1 JPEG 2000 image coding system: Core coding system, international standard edn. (2004). URL http://webstore.ansi.org
9. ISO/IEC JTC 1, SC37 Biometrics: ISO/IEC 19794-5:2005 - Biometric data interchange formats - Face Image Data, 1 edn. (2005). URL http://webstore.ansi.org International Standard

10. Aamva national standard for the driver license/identification card (2000). AAMVA DL/ID-2000
11. Wilson, C., Grother, P., Chandramouli, R.: Nist special publication 800-76-1 - biometric data specification for personal identity verification. Tech. rep., National Institute of Standards and Technology (2007). URL http://csrc.nist.gov/publications/PubsSPs.html
12. INCITS M1, Biometrics: INCITS 385:2004 - Face Recognition Format for Data Interchange, 1 edn. (2004). URL http://webstore.ansi.org. American National Standard for Information Technology
13. Blanz, V., Vetter, T.: Face recognition based on fitting a 3d morphable model. IEEE Trans. Pattern Analysis and Machine Intelligence **25**(9), 1063–1074 (2003)
14. Xiao, J., Baker, S., Matthews, I., Kanade, T.: Real-time combined 2d+3d active appearance models. In: Proc. International Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 535–542 (2004)

Face Image Quality Assessment Software

Face image quality assessment software provides multiple measurements of face image quality and determines automatically whether submitted face images are of adequate quality for a particular application.

► Photography for Face Image Data

Face Image Synthesis

► Face Sample Synthesis

Face Localization

► Face Detection

Face Matching

► Face Alignment

Face Misalignment Problem

Shiguang Shan[1], Xilin Chen[1], Wen Gao[1,2]
[1]Institute of Computing Technology, Chinese Academy of Sciences, Beijing, Peoples Republic of China
[2]Peking University, Beijing, Peoples Republic of China

Synonyms

Curse of misalignment; Face alignment error; Localization inaccuracy

Definition

The face misalignment problem, or curse of misalignment, means abrupt degradation of recognition performance due to possible inaccuracy in automatic localization of ► facial landmarks (such as the ► eye centers) in the face recognition process. Because these landmarks are generally used for aligning faces, inaccurate landmark positions imply incorrect semantic alignment between the faces or features, which can further result in matching or classification errors. Since perfect alignment is often very difficult, face recognition should be misalignment-robust, i.e., it should work well even if the landmarks are inaccurately located. To achieve this, there are three possible solutions: misalignment-invariant features, misalignment modeling, and alignment retuning.

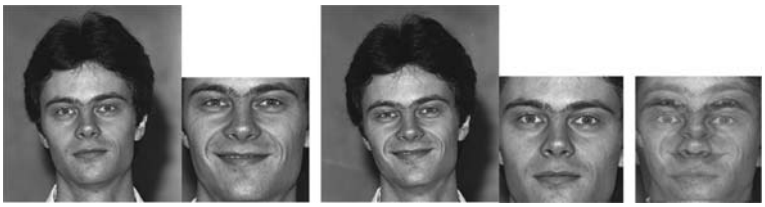
Introduction

In face recognition, before extracting features from a face image, it must be aligned properly with either the reference faces or a pre-defined general face model, with the help of some landmarks. For instance, the eye centers are generally used as control points to align all the facial images, i.e., all the faces are geometrically normalized by fixing the eye centers. Intuitively, the goal of alignment is to build the semantic correspondence among different face samples. Accurate alignment is evidently very important since the similarity (or distance) measurements generally assume the same semantics for the same feature index. However, in case of inaccurate landmark localization, this semantic alignment is broken, i.e., the component of face features extracted from the same subject with the same feature

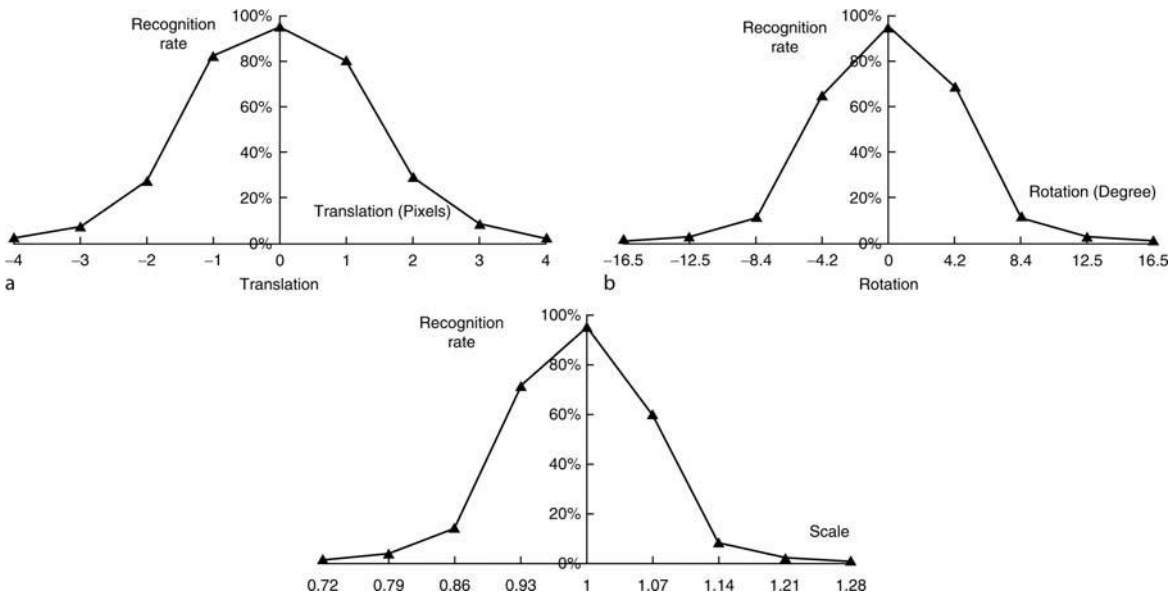
index might imply different semantics. For instance, as shown in Fig. 1, if the eyes are inaccurately localized when testing, say confused with the eyebrows, it will result in ridiculous matching eyes with eyebrows, possibly also matching nose with mouth, which will evidently lead to an incorrect classification. The above-mentioned misalignment is actually equivalent to affine transformation, i.e., translation, scaling, and rotation.

To demonstrate how much misalignment can degrade the face recognition systems [1], experiments were conducted on the FERET face database to evaluate the performance variance of Fisherfaces method [2] against the degree of misalignment. The results are shown in Fig. 2a–c, how the rank-1 recognition rates change with the misalignment degree in translation, rotation, and scale respectively. It is clear that the rank-1 recognition rate of the Fisherfaces method degrades abruptly with the increase in the misalignment. For example, 10% decrease is observed for misalignment due to a pixel translation, while 20% for misalignment of 4.2° of rotation, and almost 30% for 0.07 scale changing. Such abrupt degradation of the performance is hardly acceptable for a practical face recognition system in which misalignment of one or two pixels is almost unavoidable. Therefore, it is a problem that needs more attention.

For face recognition, aligning faces only according to the eye centers imply much more than simple affine transformation in case other variations are presented such as pose and expression. As shown in Fig. 3, the eye centers are aligned perfectly, but other features are not aligned correctly due to the 3D rotation of the head.



Face Misalignment Problem. **Figure 1** Example of misalignment caused by incorrect facial landmarks. The rightmost image is the blending of the two misaligned images, from which one can imagine how much misalignment can affect the effective matching of two biometric traits.



Face Misalignment Problem. **Figure 2** The rank-1 recognition rates of Fisherfaces against the degree of misalignment in translation, rotation and scale [1].

Face Misalignment Problem. Figure 3 Example of misalignment caused by pose variation. The rightmost image is the blending of the two misaligned images, from which one can see much misalignment in nose, mouth, and chin area, though the eye centers have been aligned correctly.

Possible Solutions

Since misalignment problem results from inaccurate (even incorrect) alignment, it is a natural idea to improve the accuracy of alignment. For instance, one can localize the eye centers more accurately or locate more landmarks (e.g., as in active shape models or active appearance models). However, to the experiences of previous work on face alignment, accurate alignment is indeed a great challenge. So, one might not expect perfect alignment and has to present efficient solutions for misalignment problem. Possible solutions can be divided into three categories: invariant features, misalignment-robust classifier, and alignment retuning.

Misalignment-invariant feature based methods expect to extract from the misaligned face images "good" representations robust to the misalignment, i.e., features change little or they even do not vary with misalignment. Some filters, such as Fourier transform,can be used for this purpose, since Fourier transform is shift and rotation invariant. Gabor wavelet is also a good choice due to its locality, which has been discussed in [3]. Recently, histogram-based object representation like, histogram of Local Binary Patterns (LBP) [4] or Local Gabor Binary Patterns (LGBP) [5] are also invariant to translation and rotation, so they can be adopted as misalignment-robust features. In addition, misalignment-invariant features can also be extracted by discriminant analysis, in which misalignment is treated as within-class variation [1].

The second category of the solution tries to design misalignment-robust classifier. In [6], the authors propose to augment the gallery by perturbation and modeled the augmented gallery by Gaussian Mixture Models (GMM). In [7], the authors propose a misalignment-robust subspace learning method for face recognition, which can infer both the well-aligned face component and the misalignment parameters.

Since the problem results from incorrect alignment, the third method naturally retunes the alignment further. Note that, these methods should be clearly different from the preceding alignment algorithms in that the retuning should make use of the feedback information from the matching or classification procedure.

Related Entries

- ▶ Face alignment
- ▶ Face descriptors
- ▶ Face localization
- ▶ Feature extraction

References

1. Shan, S., Chang. Y., Gao, W., Cao, B.: Curse of mis-alignment in face recognition: Problem and a novel mis-alignment learning solution, In: Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 17–19 May 2004, pp. 314–320 (2004)
2. Belhumeur, P.N., Hespanha, J.P., et al.:: Eigenfaces vs Fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **20**(7), 711–720 (1997)
3. Shan, S., Gao, W., Chang, Y., Cao, B., Yang, P.: Review the strength of Gabor features for face recognition from the angle of its robustness to mis-alignment, In: Proceedings of 17th International Conference on Pattern Recognition (ICPR2004), Cambridge, UK, 23–26 Aug 2004, vo1. I, pp. 338–341 (2004)
4. Ahonen, T., Hadid, A., Pietikainen, M.: Face Recognition with Local Binary Patterns, In: eighth European Conference on Computer Vision, Prague, Czech Republic, May 2004, pp. 469–481 (2004)
5. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local Gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition, In: Tenth IEEE International Conference on Computer Vision, Beijing, China, 17–20 Oct 2005, pp. 786–791 (2005)

6. Martinez, A.M.: Recognizing Imprecisely Localized, Partially Occluded and Expression Variant Faces from a Single Sample per Class, IEEE Trans. Pattern. Anal. Mach. Intell. **24**(6), 748–763 (2002)
7. Wang, H., Yan, S., Huang, T., Liu, J., Tang, X.: Misalignment-Robust Face Recognition. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) 2008. Alaska, USA, 24–26 June (2008)

Face Photograph

► Photography for Face Image Data

Face Pose Analysis

Ioannis Patras
Queen Mary, University of London,
E1 4NS, London, UK

Synonyms

Face pose estimation; Face pose recognition; Head pose analysis

Definition

Face pose analysis is the process of determining the location and the orientation of a face (► Yaw/Pitch/Roll) with respect to the camera/sensor's coordination system, and the subsequent facial analysis based on that information. A typical face pose analysis system determines the head pose by analyzing the information that is contained in the facial area (typically determined by a face detection system) using models of face geometry (i.e., models of the relative location of facial landmarks such as the nose tip and the eye corners) and/or models of face appearance (i.e., models of the intensity/color variation across a face image).

Introduction

A wide variety of systems requires the reliable analysis of facial information based on the analysis of images or image sequences. The purpose of such systems is to analyze and interpret the information that is conveyed in the facial images, such as identity information or facial expression. Examples of applications are face recognition for security/surveillance (e.g., access control in buildings or airports), multimedia indexing and retrieval (e.g., searching for family photos based on who appears in them), and facial expression analysis [1] (e.g., for deception detection or for emotion recognition).

Traditionally, facial analysis assumed that the images were obtained in controlled conditions or were manually processed (e.g., cropped, resized, and rotated) such that the faces appear at the same orientation and size (e.g., the case in passport photos). However, for a large number of applications, such as face recognition in open spaces (e.g., an airport, or a tube station), it is practically impossible to introduce such controlled conditions or manually process the data in real time. In other applications, such as facial analysis for multimedia indexing and retrieval, imposing restrictions on the head pose is undesirable. Further, for applications such as in human–computer interaction the facial pose is by itself a primary source of information, and therefore, it does not make sense to restrict it. An example is a system for communication with a computer based on head gestures (such as nodding), or gaze.

Facial pose analysis addresses the needs of such applications by automatically recovering the head pose and by allowing the extraction of features that are tailored for the further analysis of facial images under the specific pose. As the size of the facial image is assumed to be normalized (i.e., cropped and resized) by the face detection module, head pose estimation typically reduces to the estimation of the three angles that specify the rotation of the head around its three axes. Of these, a distinction should be made between the estimation of in-plane rotations [i.e., head tilting (assuming a camera facing the subject)] and out-of-plane rotations caused by gestures such as head nodding or left–right head turning (assuming a camera facing the subject) (Fig. 1). The estimation of out-of-plane rotations is arguably more difficult as it involves 3D geometric transformations, and many works in face pose analysis are focused on this problem alone.

The estimation of the head pose allows the extraction of features that are tailored for further analysis of facial images under the specific pose. For this reason, facial pose analysis precedes (or overlaps with) many

Face Pose Analysis. **Figure 1** Examples of a images from a face pose database with out-of-plane rotations, that is yaw (*horizontal axis) and pitch* (vertical axis).

other facial analysis modules such as face recognition and facial expression analysis. Further, face pose estimation requires that the facial area is reasonably well localized/detected, and therefore a face detector usually precedes it. Clearly, this require face detectors that are capable of detecting faces at various poses (e.g., [2]). As pose-specific analysis can make more robust the face localization, some face detectors and face trackers [3] (i.e., modules that localize a face in the subsequent frames of a video) perform an internal coarse or a more precise pose estimation [4].

Face Pose Estimation

The core of face pose analysis is the estimation of the face/head pose from a 2D image that depicts it. This is an instance of the more general problem of estimating the 3D rotations and translations of an (potentially deformable) object from 2D images. The developed methods can be classified into two broad categories. Appearance-based methods (e.g., [5]) that rely on models of how faces appear from different viewpoints (i.e., at different poses) and geometry-based methods that rely on the localization of facial landmarks (such as eyes and nose) on the image and 3D models of the face geometry. While appearance-based methods consider the information on the whole of the facial image at once (i.e., they are global methods), geometry-based methods estimate the head pose from the information on the 2D location of parts of the face.

A typical appearance-based method transforms the facial image into a feature set that represents the image in question in a way that it allows an easy determination of the face pose, i.e., it transforms the images from a general pixel/intensity-based representation to a pose-based representation (often called view-based representation [6]). This transformation [7, 8, 9] is typically learned from (large) databases that contain facial images of individuals at different poses. Such a transformation aims to provide a representation (i.e., a feature set) in which it is easy to distinguish between variations in the appearance due to factors other than the facial pose (e.g., identity and illumination)

and variations due to each of the pose parameters (i.e., the three rotation angles). Once a face region is transformed in this way, it is easier to disregard the variations due to other sources and recover the face pose. The variability in the feature sets extracted from images that depict faces at the same pose is called intra-class variation while the variability in the feature sets of images that depict faces at different poses is called inter-class variation. A useful transform is the one that leads to feature sets that exhibit small intra-class variation and large inter-class variation.

Once the transform is learned and each facial image is transformed to a feature set, a classifier that classifies each facepose representation to a facial pose is learned [5]. Learning a transformation and a classification scheme allows the determination of the pose of a face depicted in a previously unseen (i.e., new) image. For the new image, first the face region is detected by a face detector; then, a feature set is extracted (using the learned transform), and subsequently the head pose is determined (using the learned classifier). All classifiers require the existence of a database that is used for training and which contains a set of face images for each of which the true face pose is known. In one of the simplest classifiers the pose of the face in a new image is classified to be the pose of its nearest neighbor in the database. The term nearest neighbor, we refer to the image in the database whose representation (obtained with the learned transform) is most similar to the representation (obtained with the learned transform) of the image in question.

A typical geometry-based method relies on a 3D model of the face and on establishing of correspondences between the points in the 3D face model and the points in a 2D image that depicts the face. The estimation is based on the fact that if the pose of the face, that is the location and rotations in the 3D space of the 3D face model, were known, and the ▶ camera model and camera parameters were given, then the location of any point of the 3D model (e.g., a facial landmark such as the left eye corner) on the 2D image could be calculated. Inversely, once the locations of facial landmarks on the 2D image are detected, the 3D facial pose can be estimated.

A 3D face model approximates the 3D shape of the face at a certain level of accuracy. Commonly used 3D face models (see Fig. 2) range from simple shapes such as half-a-cylinder [3, 10] or a plane [11, 12], to elaborate 3D meshes [4] that can be either generic or person-specific. As the face model is an approximation of the true face shape and the facial landmarks are typically not perfectly detected on the 2D image (e.g., due to occlusions and illumination changes), the projection of the 3D facial points on the 2D image does not coincide perfectly with the detected landmarks. For this reason, the estimation of the face pose is usually posed as an optimization problem in which the discrepancy between the expected 2D locations of the facial landmarks (as predicted by the face and the camera models) and the locations of the detected landmarks is minimized with respect to the pose parameters. In other words, during optimization we seek the pose parameters that minimize the error. As a pose transform is a rigid transform, the estimation of the pose parameters should rely only on stable facial points (such as the nose tip or the eye corners), that is, points whose location does not change with the facial expressions (such as the corner of the mouth).

Associated with the 3D geometrical model is the appearance information, that is information on how an area around a landmark is expected to appear on a 2D image. Such information is often provided in the form of a texture map (e.g., Fig. 3). Typically,

Face Pose Analysis. **Figure 2** Examples of 3D face models.

Face Pose Analysis. **Figure 3** Texture map projected on a cylindrical face model under different poses.

correspondences are established between the facial landmarks on the texture map and points on the 2D facial image using similarity measures on the appearance of small patches around them. As the reliability of the correspondences declines for small patches, geometry-based methods typically work with images of higher resolution than appearance-based methods.

It is often the case that other sources of information can be used in order to perform pose analysis. Often, the face pose needs to be estimated not in a single image but in an image sequence (i.e., face pose tracking [4]). If the frame rate is high enough, then the face pose changes sightly and smoothly from frame to frame. This prior knowledge can be incorporated in pose estimation algorithms by using various filtering techniques, with the effect that the estimated poses vary also smoothly from fame to frame. Another sources of information that is commonly used is depth information, which is obtained either from a stereoscopic camera, or from range sensors [13]. In the first case, the facial landmarks need to be localized in both the images of a stereoscopic pair [14], and from this their location in 3D space is determined. In the second case, facial landmarks need to be localized on the range data itself. In both cases, depth information provides an additional constraint on the location and pose of the 3D model of the face. Finally, infrared imaging technologies, for single or multiple sensors can also be used.

Performance Evaluation

The main challenge in face pose analysis is the fact that facial images in the same pose appear differently due to a number of factors. The most important of these factors are identity (differences in the facial characteristics between different individuals), illumination (i.e., the ambient light), occlusions (due to facial hair or other objects such as hands or glasses), and facial expressions (e.g., frowning or smiling). Such variations in appearance lead to variations in the feature set that are extracted by appearance-based methods and make difficult the correct classification. Similarly, variations in appearance make the establishment of correspondences between the texture map and the image patches in geometry-based methods difficult.

The evaluation of the performance of the face pose estimation methods is done on a collection of images that depicts faces whose poses are known, that is on an *annotated* face pose database. The term *annotated* refers to the fact that each image in the database is stored together with the corresponding "correct" pose (often called "ground truth"). The ground truth information about the pose is usually extracted during the recording of the image. An accurate method for doing so is to use additional sensors, for example, attach a magnetic sensor on the top of the head of the person, which delivers accurate pose information. Another method, which is however less accurate, is to ask the individuals to look at a certain location while the image is recorded. A third method is to manually annotate a number of stable facial landmarks (i.e., points whose location do not change with facial expressions) and use geometry-based methods to estimate the pose.

The set of images (and the corresponding poses) contained in an annotated database used for evaluation is called the *test set*. Appearance, based methods also require the existence of an annotated database that is used for learning the transform (that given an image extracts a feature set that represents it) and the classifier (that given a feature set determines the face pose). The set of images (and the corresponding poses) are contained in such a database is called the *training set*.

During the evaluation, the pose of each image contained in the test set is estimated and the difference with the ground truth pose (i.e., the error) is calculated. Usually, the average value of the error and its variance from the average value are reported. Useful estimators are the ones that have zero mean error (*unbiased*) and small variance (i.e., small spread around the average value).

Applications

Faces that are captured by cameras or other sensors in uncontrolled environments are rarely in upright position, facing the camera/sensor from a fixed distance. Images captured by surveillance cameras, in commercial films, in family photos or homemade videos, and even images captured from web cameras attached to a laptop rarely depict faces in the same pose. Therefore, face pose analysis is an integral part of all applications that require face analysis in uncontrolled environments. Imposing restrictions on the recording conditions is very often unnatural, impractical, or infeasible. In addition, head pose estimation has by itself a number of applications, for example in human–computer interaction.

The applications of face pose analysis can be divided into three main categories:

1. *Security applications in uncontrolled environments.* In applications, such as surveillance in open spaces (e.g., airports or tube stations), the question, "is this individual in the list of suspects?" or "in which other tube stations has this individual been today?" often arise and require working with facial images in arbitrary poses. Further, applications such as access control for computer login, give an extra degree of easiness if it can allow (smaller or larger) pose variations.
2. *Multimedia indexing and retrieval.* A very large portion of produced visual material, such as images in the web, films, homemade videos, and photos, depict faces. Organizing such a material according to who is depicted allows semantic access to it, that is, allows queries such as "find photos of me with my sister".
3. *Human computer interaction and behavioral analysis.* Face/head pose can be used for communication with a computer (e.g., by head nodding or as an essential step toward gaze tracking), especially in case that disabilities prohibit the use of other primary modalities such as speech. Further, the face pose and its dynamics contain information on the emotional and affective state of individuals, and therefore can be used for automatic behavioral analysis.

Summary

Recent technological advances in image (sequence) acquisition, storage and transmission, such as the development of cheap cameras and hard disks, as well as the availability of computing resources have contributed to the integration of imaging technology in our everyday lives. As face analysis moves from controlled environments to environments in which the viewpoint cannot be controlled, or applications in which the face orientation naturally changes, head pose analysis becomes an essential part of the developed systems.

Related Entries

- ▶ Face Alignment
- ▶ Face Expression Recognition
- ▶ Face Localization
- ▶ Face Tracking
- ▶ Feature Extraction

References

1. Pantic, M., Rothkrantz, L.J.M.: Automatic analysis of facial expressions: The state of the art. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12), 1424–1445 (2000)
2. Sung, K.K., Poggio, T.: Example-based learning for view-based human face detection. IEEE Trans. Pattern Anal. Mach. Intell. **20**(1), 39–51 (1998)
3. Cascia, M.L., Sclaroff, S., Athitsos, V.: Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3d models. IEEE Trans. Pattern Anal. Mach. Intell. **22**(4), 322–336 (2000)
4. Vacchetti, L., Lepetit, V., Fua, P.: Stable real-time 3d tracking using online and offline information. IEEE Trans. Pattern Anal. Mach. Intell. **26**(10), 1385–1391 (2004)
5. Li, S.Z., Lu, X., Hou, X., Peng, X., Cheng, Q.: Learning multiview face subspaces and facial pose estimation using independent component analysis. IEEE Transactions on Image Processing **14**(6), 705–712 (2005)
6. Poggio, T.: Image representations for visual learning. Science **272**(5270), 1905–1909 (1996)
7. Kirby, M., Sirovich, L.: Application of the karhunen-loeve procedure for the characterization of human faces. IEEE Trans. Pattern Anal. Mach. Intell. **12**(1), 103–108 (1990)
8. Moghaddam, B., Pentland, A.: Probabilistic visual learning for object representation. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 696–710 (1997)
9. Martínez, A.M., Kak, A.C.: PCA versus LDA. IEEE Trans. Pattern Anal. Mach. Intell. **23**(2), 228–233 (2001)
10. Xiao, J., Kanade, T., Cohn, J.F.: Robust full-motion recovery of head by dynamic templates and re-registration techniques. In: Int'l Conf. Face and Gesture Recognition, pp. 163–169 (2002)
11. Horprasert, T., Yacoob, Y., Davis, L.S.: Computing 3-d head orientation from a monocular image sequence. In: Face and Gesture Recognition, pp. 242–247 (1996)
12. Gee, A.H., Cipolla, R.: Fast visual tracking by temporal consensus. Image Vision Comput. **14**(2), 105–114 (1996)

13. Malassiotis, S., Strintzis, M.G.: Robust real-time 3D head pose estimation from range data. Pattern Recognition **38**(8), 1153–1165 (2005)
14. Pogalin, E., Redert, A., Patras, I., Hendriks, E.A.: Gaze tracking by using factorized likelihoods particle filtering and stereo vision. In: Int'l Symposium on 3D Data Processing, Visualization and Transmission, North Carolina, Chapel Hill, USA pp. 57–64 (2006)

Face Pose Estimation

▶ Face Pose Analysis

Face Pose Recognition

▶ Face Pose Analysis

Face Processing

Face processing is a term introduced at the first international workshop on face processing in video (FPiV'04) to describe image processing tasks related to extraction and manipulation of information about human faces. The most common of these tasks are face segmentation, face detection, face tracking, face modeling, eye localization, face reconstruction, face quality and resolution improvement, best face shot selection, face classification, facial expression recognition, face memorization, and face identification.

▶ Face Databases and Evaluation

Face Recognition

▶ Liveness Assurance in Face Authentication
▶ Forensic Evidence of Face

Face Recognition From Image Sequences

▶ Face Recognition, Video-based

Face Recognition in Near-Infrared Spectrum

▶ Face Recognition, Near-infrared

Face Recognition Performance Evaluation

▶ Face Databases and Evaluation

Face Recognition Using Local Features

▶ Face Recognition, Component-Based

Face Recognition, 3D-Based

Ioannis A. Kakadiaris, Georgios Passalis, George Toderici, Takis Perakis, Theoharis Theoharis
Department of Computer Science, ECE and Biomedical Engineering, University of Houston, Houston, TX, USA

Definition

Face recognition is the procedure of recognizing an individual from their facial attributes or features and belongs to the class of biometrics recognition methods. *3D face recognition* is a method of face recognition that

exploits the 3D geometric information of the human face. It employs data from 3D sensors that capture information about the shape of a face. Recognition is based on matching metadata extracted from the 3D shapes of faces. In an *identification* scenario, the matching is one-to-many, in the sense that a probe is matched against all of the gallery data to find the best match above some threshold. In an *authentication* scenario, the matching is one-to-one, in the sense that the probe is matched against the gallery entry for a claimed identity, and the claimed identity is taken to be authenticated if the quality of match exceeds some threshold. 3D face recognition has the potential to achieve better accuracy than its 2D counterpart by utilizing features that are not sensitive in lighting conditions, head orientation, differing facial expressions, and make-up.

Introduction

In recent years, among the many biometric modalities, the face has received the most interest. Not only is face recognition one of the most widely accepted modalities, but advances in processing power have allowed the development of more complex algorithms while still providing a rapid response to queries. Face recognition requires no contact with the subject, thus being more easily accepted by the public compared to other biometrics such as fingerprints.

Face recognition has been traditionally performed using 2D (visible spectrum) images, while hybrid approaches have used infrared images and 3D geometry. Infrared face recognition has not been widely adopted due to the high cost of infrared cameras necessary to acquire data. In contrast, the cost of 3D scanners has dropped significantly, so it has become feasible to deploy them in the field, and therefore, the interest in developing algorithms that use 3D data has increased.

The main reason for using information from 3D data as a biometric is that the data acquired by 3D acquisition devices are invariant to pose and lighting conditions, these being the major challenges with which face recognition algorithms must cope. Moreover, image-based face recognition algorithms are more susceptible to impostors. Indeed, an impostor may use a printout of an image of a subject allowed to enter a facility in order to break in. To avoid this, the face recognition algorithm must be coupled with liveness test algorithms. Attempting such an attack on a system based on 3D data would be much more difficult, since the attackers would need to obtain an accurate 3D model (sculpture) of the person whom they would like to impersonate.

The challenges of a 3D face recognition system are the following:

- *Accuracy gain*: A significant gain in accuracy with respect to 2D face recognition systems must justify the introduction of 3D recognition systems.
- *Efficiency*: 3D capture devices generate substantially more information than 2D cameras. Using this large volume of information is expensive in terms of computation time and storage requirements. Therefore, the algorithms developed need to be efficient both in time and space, by using the appropriate metadata.
- *Automation*: The system must be completely automated. It is therefore not acceptable to assume user intervention, such as for the location of key landmarks in a 3D facial scan.
- *Capture devices*: 3D capture devices were mostly developed for medical and other low-volume applications and suffer from a number of drawbacks, including artifacts, small depth of field, long acquisition time, multiple types of output, and high price. A deployable 3D face recognition system must be able to process several persons a minute, if it is to be used in high-traffic areas.
- *Testing databases*: There are few large databases of 3D faces which are widely accepted for objectively testing the performance of 3D face recognition systems. More such databases are needed to ensure proper testing of the system.
- *Robustness*: The system must perform robustly and reliably under a variety of conditions (e.g., lighting, pose variation, facial feature variation).

An Integrated 3D Face Recognition System

The authors have developed a fully automatic system [1] which is capable of using 3D data as input, along with a facial model, to output metadata information. The metadata are then used for recognition. The facial model has been constructed only once, and it can handle objects belonging to the same class (i.e., faces). Once the data are acquired, the model is fitted to the data

and used to generate a geometry image and a normal map, which are transformed into the wavelet domain. Only a small number of the wavelet coefficients are stored as metadata and used for comparison.

Our recognition procedure can be divided into two distinct phases: *enrollment* and *recognition*.

Enrollment: Raw data acquired by the 3D scanner are converted to metadata and stored in a database (gallery). The following steps describe the conversion from raw data to metadata (Fig. 1):

1. *Acquisition*: The sensor acquires raw data which are converted into a polygonal representation. A preprocessing step takes place to alleviate scanner-specific issues.
2. *Alignment*: The data are aligned into a unified coordinate system using a multi-stage alignment method.
3. *Deformable model fitting*: The ▶ annotated face model (AFM) is fitted to the data.
4. *Metadata generation*: Geometry and normal map images are derived from the fitted model and wavelet analysis is applied to extract a reduced set of the most significant coefficients.

Recognition: Metadata extracted from a face probe (using the same steps as for enrollment) are directly compared with metadata retrieved from the database gallery using a distance metric.

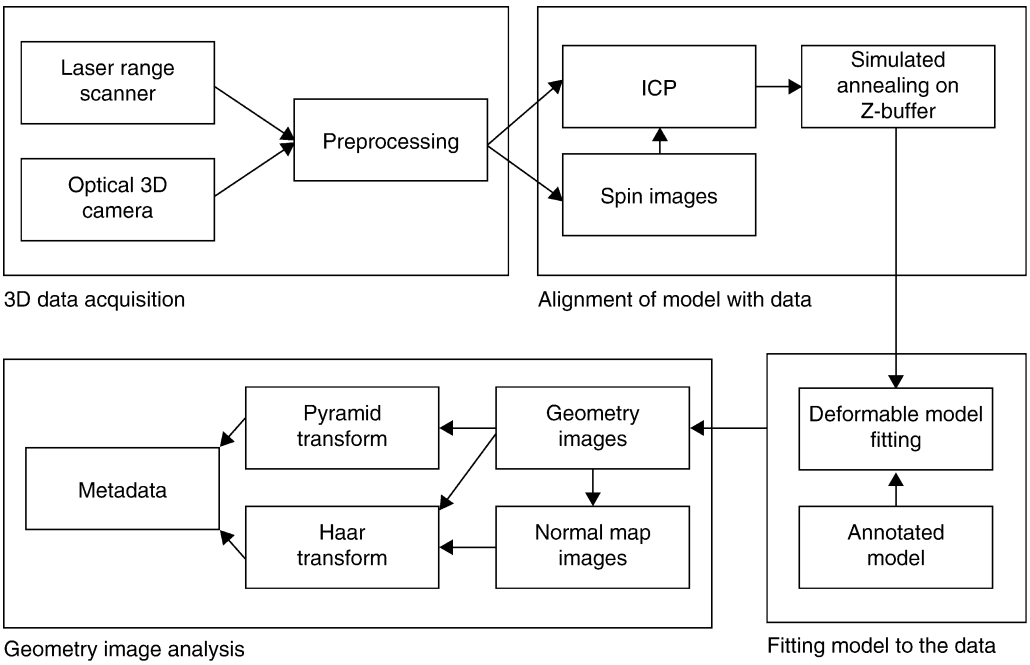
Data Acquisition and Preprocessing

In general, the current generation of scanners outputs either a range image or 3D polygonal data. The purpose of preprocessing the data is the elimination of any sensor-specific issues and the unification of data from different sources into a common format (Fig. 2). The preprocessing consists of the following filters that operate on both the native representations and on 1-neighbors, and are applied in the given order:

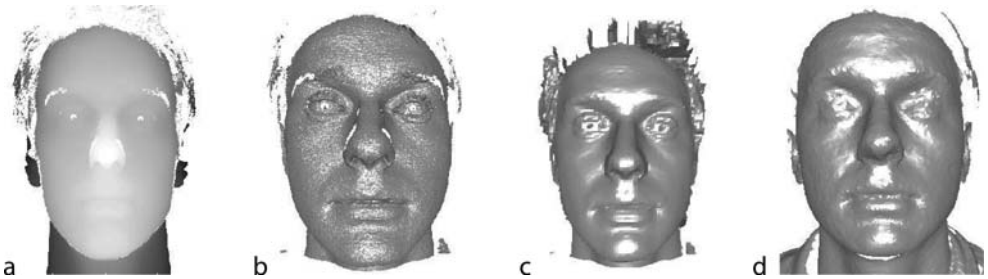
- *Median cut*: This filter removes spikes from data acquired by using laser scanners.
- *Hole filling*: Eliminates holes produced by laser scanners in certain areas such as eyes and eye brows.
- *Smoothing*: A smoothing filter is applied to remove white noise.
- *Subsampling*: The deformable model fitting effectively resamples the data, making the method less sensitive to data resolution without losing performance in the recognition phase. Subsampling further reduces the noise in the geometry.

Annotated Face Model

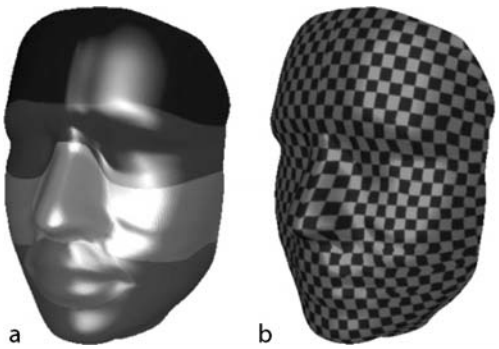
Our approach introduces an AFM, which is constructed only once and is used in the alignment, fitting,



Face Recognition, 3D-Based. **Figure 1** Enrollment phase of the proposed integrated 3D face recognition system.



Face Recognition, 3D-Based. Figure 2 Sensor-dependent preprocessing. Laser range scanner: (**a**) input depth image, (**b**) raw polygonal data (200,000 triangles), and (**c**) processed data (16K). Stereo camera: (**d**) raw data (34,000 triangles).



Face Recognition, 3D-Based. Figure 3 AFM: (**a**) Annotated facial areas and (**b**) texture used to demonstrate parameterization.

and metadata generation [1]. The model is anthropometrically correct according to Farkas' work [2], and is annotated into different facial areas (e.g., mouth, nose, eyes) (Fig. 3). Applying a continuous global UV parameterization on the model, all vertices of the model from R^3 to R^2 and vice versa have been mapped. Therefore, the model is defined both as polygonal data in R^3 and as a geometry image in R^2 [1, 3].

A ▶ geometry image is a regular sampling of the model represented as a 2D image with three channels, each channel corresponding to the *x*, *y*, and *z* coordinates of the 3D object. Since local neighborhoods on the mesh are preserved (i.e., neighboring vertices are preserved even in the geometry image), an approximated version of the original mesh can be reconstructed from the geometry image. The number of channels in the geometry image can be greater than three, as apart from geometric information, texture and annotation can also be encoded.

Alignment

Our work on face recognition has indicated that alignment (pose correction) is a key part of any geometric approach. So, before fitting, align each preprocessed dataset with the AFM. The alignment stage computes a rigid transform, combining rotation and translation, which brings the data as close as possible to the model and is robust and accurate even when relatively large deformations (facial expressions) occur in the input data. Our alignment algorithm is a multi-stage algorithm which propagates the alignment variables from one stage to the next [1]. The first algorithm is more resilient to local minima, while the next two algorithms provide greater alignment accuracy:

- *Spin images*: The purpose of the first step is to establish a plausible initial correspondence between the model and the data. This step can be omitted if the arbitrary rotations and translations in the databases are not expected. A spin image is a representation of the geometric neighborhood around a specific point [4]. To register two shapes, the correspondences between the individual spin images must be found. These correspondences are grouped into geometrically consistent groups and the transformations they yield are verified by checking if they rotate the data by an acute angle (based on the assumption that a given face does not have an upside down pose or an opposite orientation from the camera). This check is essential due to the bilateral symmetry property of the human face.
- *Iterative closest point* (*ICP*): The main step of our alignment process uses the ICP algorithm [5] extended in a number of ways. The ICP algorithm

solves the registration problem by minimizing the distance between the two sets of points. The annotated model is exploited by assigning different weights to different face regions. Additionally, pairs containing points on surface boundaries are rejected. This ensures that no residual error is introduced into ICPs metric by the non-overlapping parts of two surfaces. Finally, if the resulting transformation is not satisfactory, the option of running the trimmed ICP algorithm [6] is available.

- *Simulated annealing on z-buffers*: This is a refinement step that ensures that the model and the data are well aligned. The idea is to refine alignment by minimizing the differences between the z-buffers of the model and data. A global optimization technique has been employed, known as enhanced simulated annealing (ESA) [7], to minimize the z-buffer difference [8]. The higher accuracy of this step can be attributed to the fact that the z-buffers effectively resample the data which results in independence from the data's triangulation.

Deformable Model Fitting

The purpose of fitting the model to the data is to capture the geometric information of the desired object. In order to fit the AFM to the raw data, a subdivision-based deformable model framework [1] is used. When the deformation concludes, the AFM acquires the shape of the raw data. This establishes a dense correspondence between the AFMs surface and the raw data's vertices. Additionally, since the deformation has not violated the properties of the original AFM, the deformed AFM can be converted to a geometry image. The extracted geometry image encodes the geometric information of the raw data (Fig. 4). Note that the deformable model framework discards data not belonging to the face and successfully handles artifacts without any special preprocessing.

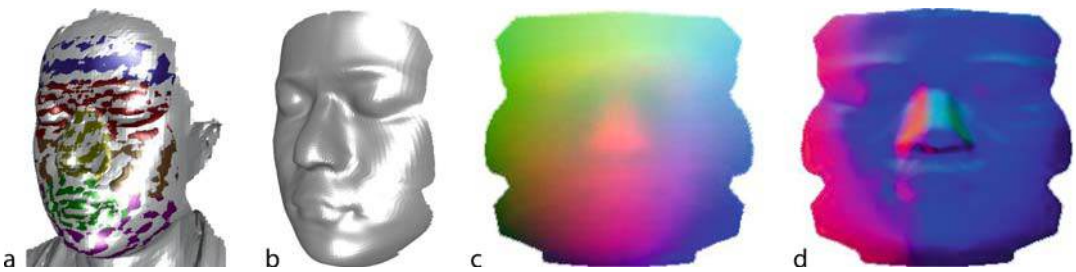
The fitting framework is an implementation of the ▶ elastically adaptive deformable models [9] using subdivision surfaces [10]. The Loop subdivision scheme [11] has been selected since it produces a limit surface with C^2 continuity, while only 1-neighborhood area information is needed for each vertex. The AFM is used as the subdivision surface's control mesh, thus determining the degrees of freedom, while the limit surface is used to solve the following equation:

$$M_q \frac{d^2 q}{dt^2} + D_q \frac{dq}{dt} + K_q q = f_q,$$

where q is the control points vector, M_q is the mass matrix, D_q is the damping matrix, K_q is the stiffness matrix, and f_q are the external forces. The external forces drive the deformation. The stiffness matrix defines the resistance against the deformation, while the mass and damping matrices control the velocity and the acceleration of the vertices. This equation is solved based on the finite element method (FEM) approximation. During this process the AFM gradually acquires the shape of the raw data.

Metadata Generation

The deformed model that is the output of the fitting process is converted to a geometry image, as depicted in Fig. 4(c). The geometry image regularly samples the deformed model's surface and encodes this information on a 2D grid. The grid resolution is correlated with the resolution of the AFMs subdivision surface. From the geometry image, a normal map image (Fig. 4(d)) is also constructed. The normal map



Face Recognition, 3D-Based. **Figure 4** Full face model after fitting: (**a**) Fitted model overlayed on the face data, (**b**) fitted model geometry, (**c**) corresponding geometry image, and (**d**) corresponding normal map.

contains the 3D normal vectors to the surface as its pixel values [1].

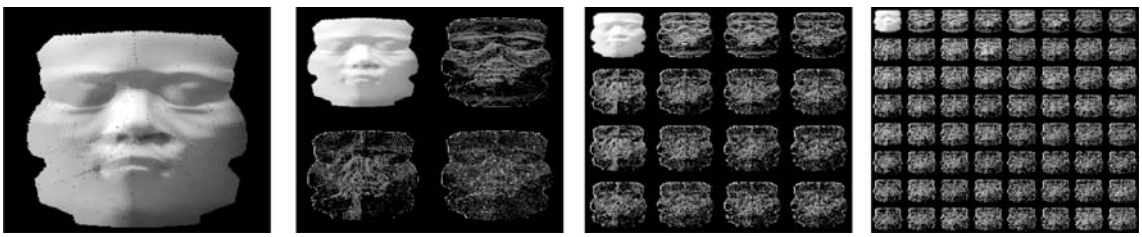
The three channels (components *X*, *Y*, and *Z*) of the normal map and geometry image have been treated as separate images. Each component is analyzed using a wavelet transform and the coefficients are stored as metadata. Two different transforms have been used, the Haar and Pyramid transforms, thus obtaining two sets of coefficients. The Pyramid transform is a more computationally intensive transform, and therefore, we may choose not to use it if the system needs to be tuned for speed. The Haar transform is applied on both the normal map and the geometry image, while the Pyramid transform is applied only on the geometry image.

- *Haar wavelets*: The choice of Haar wavelets was based on their properties. The transform is conceptually simple and computationally efficient. The Haar wavelet transform is performed by applying a low-pass filter and a high-pass filter on a one-dimensional input, and then repeating the process on the two resulting outputs. Since we are working with images, there will be four outputs for each level of the Haar wavelet (Low–Low, Low–High, High–High, High–Low). A level 4 decomposition is computed, meaning that the filters are applied four times, which yields 256 (16×16) wavelet packets (Fig. 5). Each packet contains a different amount of energy from the initial image. It is possible to ignore most of the packets without losing significant information and store the same subset of the most significant coefficients as metadata. This allows an efficient direct comparison of coefficients of two images without the need for reconstruction.
- *Pyramid transform*: The second transform decomposes the images using the complex version of the steerable pyramid transform [12], a linear, multi-scaled, multi-orientation image decomposition algorithm. The resultant representation is translation-invariant and rotation-invariant. This feature is desirable to address possible positional and rotational displacements caused by facial expressions. To maintain reasonable image resolution and computational complexity, our algorithm applies a 3-scale, 10-orientation complex steerable pyramid transform to decompose each channel of the geometry image. Only the low-pass orientation subbands at the farthest scale are stored as metadata. This enables us to compare the subband coefficients of two images directly without the overhead of reconstruction.

Distance Metrics

In the recognition phase, the comparison between two subjects (gallery and probe) is performed using the metadata information. The coefficients of the geometry image are kept as metadata, and the normal map of each dataset. Additionally, there may be two coefficient types for each: the Haar coefficients and, optionally, the Pyramid coefficients. To compare the metadata, there is a need to define a distance metric for each type of coefficient:

- *Haar metric*: In the case of Haar wavelets, the metric used is weighted L^1 on each component independently. The total distance is the sum of the distances computed on all components.
- *Pyramid metric*: A modified version of the complex version of the structural similarity index (CW-SSIM) [13] is used. CW-SSIM iteratively measures the similarity indices between two sliding windows



Face Recognition, 3D-Based. **Figure 5** Haar wavelet analysis for the normal map image: (**a**) zero level, (**b**) first level, (**c**) second level, (**d**) third level. Note that the real numbers were mapped to a gamma corrected grey-scale for visualization purposes.

placed in the same positions on the two images, and uses the weighted sum as a final similarity score.
- *Fusion*: When both types of coefficients are used, the distances given by the Haar and the Pyramid metrics are fused. A weighted sum of the two distances is used as a fusing score.

3D Face Recognition Hardware Prototype System

A field-deployable prototype system has been built and is operational at the University of Houston. It consists of a 3dMD^{TM} 3D camera (1-pod configuration) which is connected to a laptop. The color camera of the pod captures a continuous video stream which is used to detect whether a person is facing the 3D camera. When the subject is facing the camera and remains relatively still for more than 2 s, the system triggers the 3D camera and the geometry data of the individual's face are captured. Each of the cameras has a resolution of 1.2 megapixels. The entire capture process takes less than 2 ms, and it produces a mesh with less than 0.5 mm RMS error (as quoted by the manufacturer).

The system can either enroll the subject into the database or perform a scenario-specific task. In an identification scenario, the system will display the closest five datasets to the operator. In a verification scenario, the system will determine whether the subject is who he/she claims to be, based on a preset distance threshold.

The system's field-deployable characteristics are:

- *Automation*: All methods utilized are fully automated, requiring no interaction with a user. The system is capable of detecting when a subject is within range by using a face detector implementation, and initiating the enrollment or authentication procedures automatically.
- *Space efficiency*: The raw 3D data produced by most scanners are of several MB. After the enrollment phase, the system needs to keep only the metadata.
- *Time efficiency*: The enrollment phase is the most time consuming, as the time delay to convert the raw scanner data to the final metadata is 15 s. In the authentication phase of an identification or verification scenario, only the stored metadata are utilized. The system can compare the metadata of enrolled subjects at a rate of 1,000/s on a typical modern PC (3.0 GHz P4, 1 GB RAM).

Performance Evaluation

Databases

The results on 3D face recognition are reported using two databases. The first is the well known FRGC v2 database and the second is a collection of 3D faces acquired at the University of Houston (UH). To demonstrate the sensor-invariant nature of the proposed system, the UH database is combined with FRGC v2.

The *FRGC v2* database [14, 15] contains 4,007 3D scans of 466 persons. The data were acquired using a Minolta 910 laser scanner that produces range images with a resolution of 640×480. The scans were acquired in a controlled environment and contain various facial expressions (e.g., happiness or surprise). The subjects are 57% male and 43% female, with the following age distribution: 65% 18–22 years old, 18% 23–27 and 17% 28 years or over. The database contains annotation information, such as gender and type of facial expression.

The *UH* database contains 884 3D facial datasets acquired using our 3dMD^{TM} system (with 1-pod and 2-pod setups) over a period of one year. The data acquisition protocol was the following:

For each subject:

- Remove any accessories (e.g., glasses).
- Acquire a dataset with neutral expression.
- Acquire several datasets while the subject reads loudly a predefined text (thus assuming facial expressions).
- Put on the accessories and acquire a dataset with neutral expression.

The UH database is more challenging compared to the FRGC v2 as the subjects were encouraged to assume various extreme facial expressions and in some cases accessories were present. The resulting extended database contains a total of 4,891 datasets, 82% acquired using a laser scanner, 18% acquired using an optical camera, and, to the best of our knowledge, is the largest 3D facial database reported.

Performance Metrics

Two different scenarios have been employed for the experiments: *identification* and *verification*. In an identification scenario, divide the database into probe and gallery sets so that each subject in the probe set has

exactly one match in the gallery set. To achieve this, use the first dataset of every individual as gallery and the rest as probes. The performance is measured using a cumulative match characteristic (CMC) curve and the rank-one recognition rate is reported.

In the verification scenario, measure the verification rate at 0.001 false accept rate (FAR). The verification rate is defined as the fraction of datasets that are positive (e.g., claiming to be who they really are), and are classified as positive. The FAR is defined as the fraction of datasets that are negative (e.g., pretending to be somebody else), but are classified as positive. The results are plotted using a receiver operating characteristic (ROC) curve which plots verification rate as a function of FAR. The FRGC v2 database defines three possible selections of datasets (referred to as ROC I, ROC II, and ROC III). In ROC I, all the data are within semesters, in ROC II, they are within one year, while in ROC III, the samples are between semesters. These experiments are of increasing difficulty.

Experiment 1: Wavelet Transforms

The purpose of this experiment is to evaluate the performance of the two wavelet transforms, and to provide a reference score on the FRGC v2 database. Using a fusion of the two transforms, our system yielded a verification rate of 97.3% (for ROC I at 0.001 FAR), while separately for the Haar transform a rate of 97.1% and for the Pyramid transform a rate of 95.2% were achieved (Table 1).

Even though the Pyramid transform is computationally more expensive, it is outperformed by the simpler Haar wavelet transform. This can be attributed to the fact that in the current implementation, the Pyramid transform utilizes only the geometry images and not the normal map images. The fusion of the two transforms offers more descriptive power, yielding higher scores, especially in the more difficult experiments of ROC II and ROC III, as depicted in Table 1.

Face Recognition, 3D-Based. **Table 1** Verification rates of our system at 0.001 far using different transforms on the Frgc V2 database

	ROC I	ROC II	ROC III
Fusion	97.3%	97.2%	97.0%
Haar	97.1%	96.8%	96.7%
Pyramid	95.2%	94.7%	94.1%

To the best of our knowledge, this is the highest performance reported on the FRGC v2 database for the 3D modality.

Experiment 2: Facial Expressions

Facial expressions have traditionally decreased the performance of face recognition systems. In this experiment, the authors evaluate the impact of facial expressions on the performance of the system. All datasets in FRGC v2 are annotated, and one of the categories recorded is the facial expression. The authors chose to divide the database into two distinct sets: the first set contains non-neutral facial expressions only, while the second set contains datasets that were annotated as having a neutral facial expression.

The performance of the two subsets is compared to the performance on the entire set at 0.001 FAR in Table 2. The average decrease of 1.56% in verification between the full database and the subset containing only facial expressions is very modest when compared to most other systems, given the fact that this subset contains the most challenging datasets from the entire database and is fully automatic. The small decrease in performance can be attributed to the use of the deformable model framework and the AFM.

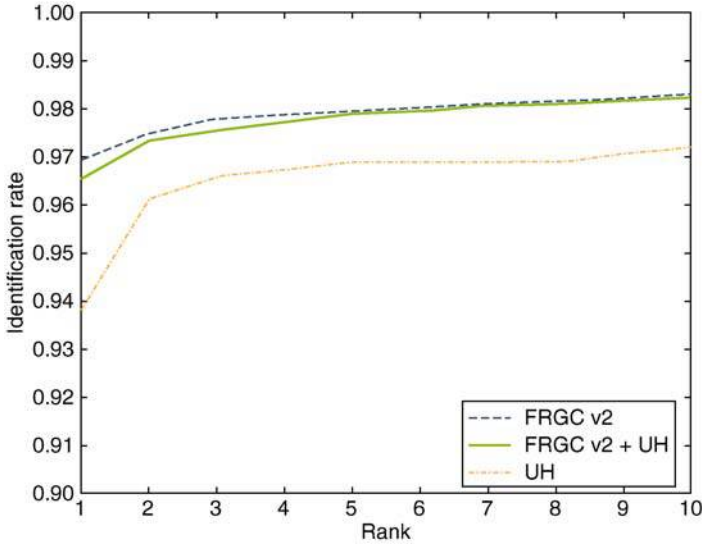
Experiment 3: Multiple Sensors

The purpose of this experiment is to evaluate the performance of our system using data from multiple sensors. Verification experiments depend heavily on the pairs of datasets chosen for evaluation. In the absence of any standard way of designing such experiments, opt for an identification experiment, which is considered to be more representative and more easily duplicated.

This identification experiment was conducted on different databases: FRGC v2 database, with 466 gallery and 3,541 probes (laser scanner), UH database with

Face Recognition, 3D-Based. **Table 2** Performance of our system at 0.001 far on the full FRGC V2 database, on a subset containing only non-neutral facial expressions and on a subset containing only neutral expressions

	ROC I	ROC II	ROC III
Full Database	97.3%	97.2%	97.0%
Non-neutral	95.6%	95.6%	95.6%
Neutral Expressions	99.0%	98.7%	98.5%



Face Recognition, 3D-Based. Figure 6 System performance for identification experiment on different databases: FRGC v2 database with 466 gallery and 3,541 probes (laser scanner), UH database with 240 gallery and 644 probes (optical scanner) and FRGC v2+UH database with 706 gallery and 4,185 probes (both scanners).

240 gallery and 644 probes (optical scanner) and FRGC v2+UH database with 706 gallery and 4,185 probes (both scanners). On the FRGC v2 dataset, the rank-one identification rate was 97.0%, while for the UH set, the system achieved 93.8%. Fig. 6 depicts the full CMC curve. The combined experiment yielded a rank-one recognition rate of 96.5%, which represents a drop in performance of only 0.5% when compared to the original FRGC v2 experiment, demonstrating the system's robustness when data from multiple sensors are included in the same database.

Conclusion

The authors presented algorithmic solutions to the majority of the challenges faced by field-deployable 3D facial recognition systems. By utilizing an annotated deformable model, the 3D geometry information is mapped onto a 2D regular grid, thus combining the descriptiveness of 3D data with the computational efficiency of 2D data. A multi-stage fully automatic alignment algorithm and the advanced wavelet analysis resulted in robust state-of-the-art performance on the publicly available FRGC v2 database. Our multiple-sensor database pushed the evaluation envelope one step further, showing that both accuracy and robustness can be achieved when data from different sensors are present, through sensor-oriented preprocessing. Proof of concept is provided by our prototype system which combines competitive accuracy with storage and time efficiency.

Related Entries

- ▶ Anatomy of Face
- ▶ Deformable Models
- ▶ Face Localization
- ▶ Face Pose Analysis
- ▶ Face Recognition: Component-based
- ▶ Face Recognition: Shape vs Appearance

References

1. Kakadiaris, I., Passalis, G., Toderici, G., Lu, Y., Karambatziakis, N., Murtuza, N., Theoharis, T.: 3D face recognition in the presence of facial expressions: an annotated deformable model approach. IEEE Trans. Pattern Anal. Mach. Intell. **29**(4), 640–649 (2007)
2. Farkas, L.: Anthropometry of the Head and Face. Raven Press, NY (1994)
3. Gu, X., Gortler, S., Hoppe, H.: Geometry images. In: Proceedings of SIGGRAPH, pp. 355–361, San Antonio, TX, USA, July (2002)
4. Johnson, A.: Spin-images: a representation for 3-D surface matching. Ph.D. Thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August (1997)

5. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
6. Chetverikov, D., Svirko, D., Stepanov, D., Krsek, P.: The trimmed iterative closest point algorithm. In: Proceedings of the International Conference on Pattern Recognition, vol. 3, pp. 545–548. Quebec City, Canada (2002)
7. Siarry, P., Berthiau, G., Durbin, F., Haussy, J.: Enhanced simulated annealing for globally minimizing functions of many-continuous variables. ACM Trans. Math. Software **23**(2), 209–228 (1997)
8. Papaioannou, G., Karabassi, E., Theoharis, T.: Reconstruction of three-dimensional objects through matching of their parts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(1), 114–124 (2002)
9. Metaxas, D., Kakadiaris, I.A.: Elastically adaptive deformable models. IEEE Trans. Pattern Anal. Mach. Intell. **24**(10), 1310–1321 (2002)
10. Mandal, C.: A dynamic framework for subdivision surfaces. Ph.D. Thesis, University of Florida (1998)
11. Loop, C.: Smooth subdivision surfaces based on triangles. M.Sc. Thesis, Department of Mathematics, University of Utah (1987)
12. Portilla, J., Simoncelli, E.P.: A parametric texture model based on joint statistic of complex wavelet coefficients. Int. J. Comput. Vis. **40**, 49–71 (2000)
13. Wang, Z., Simoncelli, E.: Translation insensitive image similarity in complex wavelet domain. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. II, pp. 573–576. Philadelphia, PA, USA (2005)
14. Phillips, P.J., Flynn, P.J., Scruggs, W.T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, vol. 1, pp. 947–954. Gaithersburg, MD, USA (2005)
15. Phillips, P.J., Scruggs, W.T., O'Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M.: FRVT 2006 and ICE 2006 Large-Scale Results. NISTIR 7408, March (2007)
16. Kirkpatrick, S., Gelatt, C., Vecchi, M.: Optimization by simulated annealing. Science **22**(4598), 671–680 (1983)

Face Recognition, Component-Based

Onur C. Hamsici, Aleix M. Martinez
The Ohio State University, Columbus, OH, USA

Synonyms

Face recognition using local features; Part-based face recognition

Definition

A major problem in face recognition is to design algorithms that are invariant to those image changes typically observed when capturing faces in real environments. A large group of important image variations can be addressed using a component-based approach, where each face is first analyzed by parts and then the results are combined to provide a global solution. The image variations that are generally tackled with this approach are those due to occlusion, expression, and pose [1]. It has been argued that these changes have a lesser effect on local regions than to the whole of face image. Differences exist on how to formulate the component-based approach. Some of the algorithms use local information and combine these using a global decision maker. Some extract the important local parts to represent the face distributions, while others learn the distribution of the components generated by the variations. A summary of these techniques is given in this essay.

Introduction

Component-based face recognition algorithms include those that use some local information of the face to do recognition of the whole. These algorithms are very popular, since the local information is generally more robust to many of the typically seen parameter variations of the face. This is especially true if one does recognition based on the texture (i.e., pixel information) of the face.

One of these parameters is the location of the fiducial points in the face. These fiducial points are necessary to align all faces with respect to one another. However, it is not usually possible to obtain the exact location of these points automatically. This generates *imprecise localizations* which will further decrease the performance of the recognition algorithms [1]. Component-based algorithms can also be made more robust to these errors of localization. This is because some of the local features may be localized more precisely than the other ones and, hence, lead to better recognition rates.

A similar advantage is also seen in expression and pose changes. In this case, some local components of the face may have less expression changes (such as the nose region when a person smiles) or maybe less affected by pose changes (such as the eye region that is in the opposite side of the head).

Moreover, brightness changes are known to be handled better when the face is represented by components. It is because the face is a nonconcave structure, resulting in different lightings across it. For example,

the right and left side of a face may be lighted with totally different lighting conditions and, hence, may lead to different pixel levels. Trying to handle these changes using a global approach may fail due to the possible lighting changes. Simple intensity normalization processes can be used to eliminate part of the lighting differences when using a component-based approach [2].

Another advantage of using component-based algorithms is the stability for the possible occlusions over the faces. Even when half of the face is occluded, as for example with a scarf or large eyeglass, a component-based algorithm can still employ the information of the other half of the face image to do recognition.

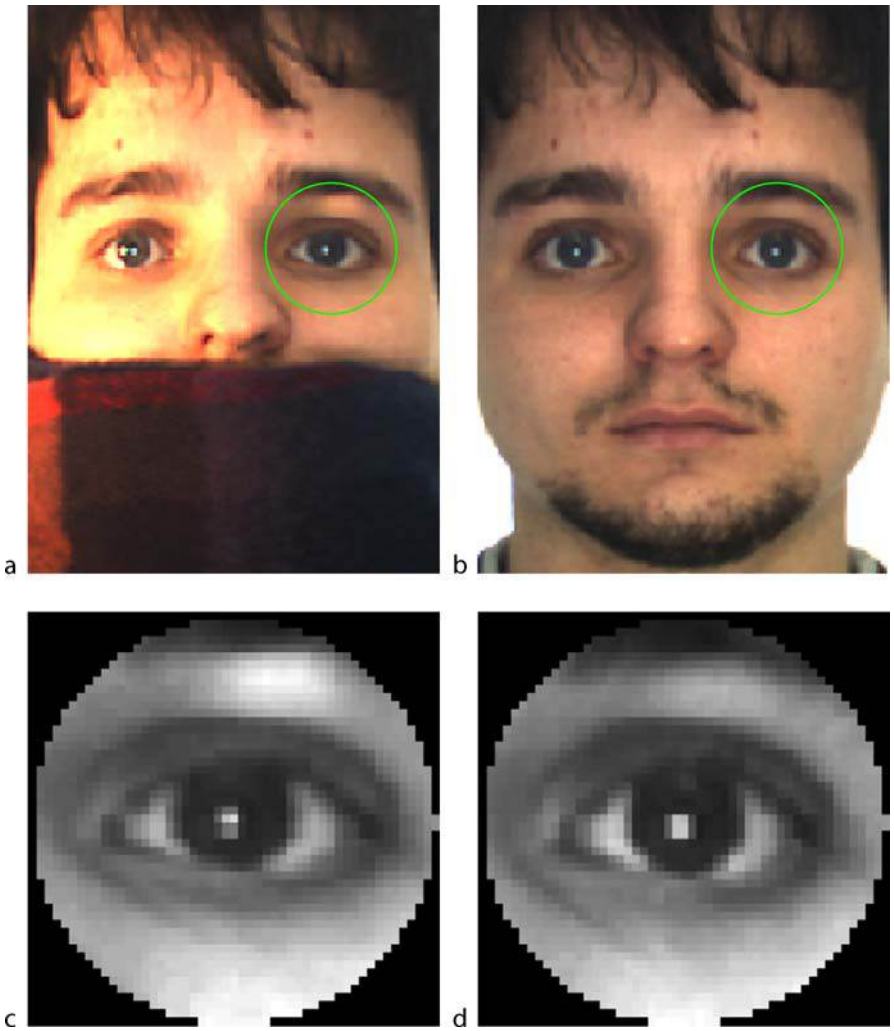
Figure 1 shows some of the advantages of the local approach representation just described. Although, there is an extreme lighting change and occlusion of the face, the local right eye regions are very similar in both images and shall lead to a successful classification.

Because of these advantages component-based face recognition algorithms are preferred approaches in many real settings. In the following sections some of the most used algorithms defined thus far have been investigated. A discussion as well as the pros and cons of each technique, have also been provided.

Component-Based Face Recognition

Component-Based Graphs

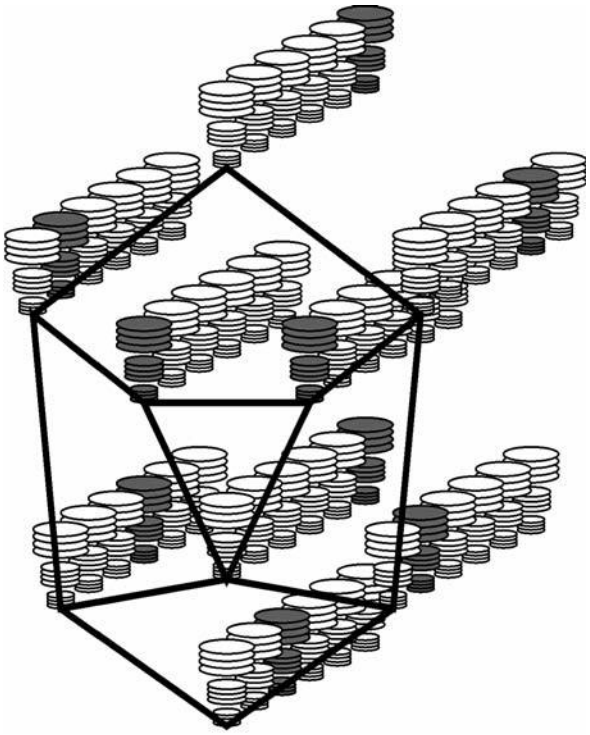
Building a system that is not skewed by localization errors seems close to impossible. This is due to the many



Face Recognition, Component-Based. Figure 1 Although the faces shown in **(a)** and **(b)** have extreme lighting changes and occlusions, the corresponding right eye regions in **(c)** and **(d)** are very similar.

additional variables that define the face image. These include pose, expression lighting, etc. Researchers have found a solution to this problem. In this solution, algorithms have been developed that consider the range of the localization errors over the given image. One solution, as defined in the section to follow, is to learn the set of textural changes due to localization errors. A robust alternative is to handle this set by designing an algorithm that depends on the local information. This is done in [3] with the *Dynamic Link Architecture* (DLA) and in [4] with the *Elastic Bunch Graph Matching* (EBGM). Both of these algorithms use the local information by dividing each image into a set of patches. While DLA extracts these patches by dividing the image with a grid structure, EBGM uses the regions around the fiducial points. Both DLA and EBGM extract features from the patches by filtering them with complex Gabor jets and using the magnitude of their outputs. The rationale behind the use of these features is grounded in the fact that ▶ Gabor jets are known to be less affected by lighting changes. In addition, EBGM uses the phase of the filtered patch to locate the nodes more accurately and to differentiate the patches that have the same magnitude.

One major difference between these two approaches is seen in the graph representation of the components. In DLA the spatial information between the patches is defined using a graph with nodes representing the grid parts of a face in the image plane. A proper matching algorithm between the images were proposed using the spatial information (hidden in edges of the graph) and the local information (hidden in vertices of the graph). In EBGM the face bunch graph (FBG) is defined over the fiducial points such that each node represents the Gabor jet outputs for several variations of a fiducial point, i.e., the node related with eyes may include an eye bunch that is closed, open, left–right pupil, and so on. Figure 2 shows an example of FBG. Here each node corresponds to a fiducial point where the set of discs are the Gabor jets related with the corresponding region. Bunch of set of discs represents the variability in the faces around that region. Matching is done using an elastic bunch graph matching algorithm which considers the size change of the FBG and location change of the nodes to optimize the graph similarity. Once the graph is obtained the recognition is done by calculating the similarity between the test image and the training image graphs. The match is found across all possible variations of Gabor jets. In Fig. 2 a possible match is shown by gray marking. This component-based strategy has proven to yield good results for frontal face recognition and reasonable results under pose variations.

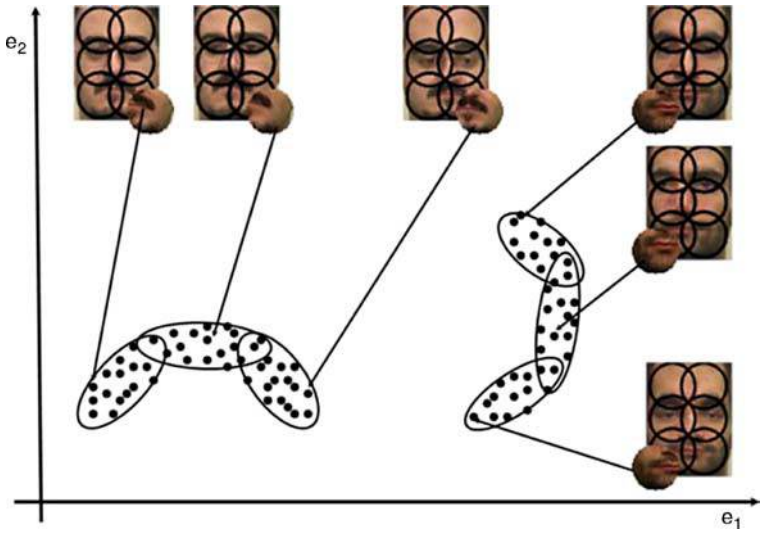


Face Recognition, Component-Based. Figure 2 A Face Bunch Graph (FBG) is shown, which represents all possible variations across faces. Each jet is represented by a stack of discs. In the matching process, the best fitting bunch of jets (shown in gray) is selected from a bunch of jets attached to a single node. © IEEE 1997.

Modeling Components

As mentioned above, some of the important problems to deal with in face recognition are imprecise localization, ▶ partial occlusion, and expression variation. Another important problem is given by the traditionally small number of training samples per class, since one usually has access to just a few images per subject. In [1] these problems are handled by means of a probabilistic approach.

▶ Imprecise localizations arise from not knowing the exact position of the fiducial points located in the face. Not only automatically detected, but also hand-marked feature locations include imprecise localizations.



Face Recognition, Component-Based. **Figure 3** This figure shows an approach used to model all possible warped face images according to the localization error of the localization algorithm. After division of each face image into K local areas, the localization error is estimated (for each of these local areas) using a mixture of Gaussians. © IVC 2006.

Although the additional noise may be small in the image plane, the corresponding images in the Euclidean space may deviate from their classes considerably. To handle this problem, [1] proposes to model the localization error by synthetically generating images that may be observed under a given error. Then, these images are modeled by means of a mixture of Gaussian distribution. With this approach one extends on the original set of images to a larger one that most appropriately represents image variations. Figure 3 shows this for a face image division into six local parts. All possible images of a local region, which are generated by localization error, are modeled using a mixture of Gaussians.

Another major problem is partial occlusions. This is handled by dividing the face into a set of independent regions. This allows us to avoid using nonface areas that may distract the recognition process in a global approach. The component-based approach described above has shown to be superior to global techniques and voting strategies in [5] using a large set of images extracted from the AR database [6].

Expression changes are eliminated in a similar manner. In this case, a weighting scheme is used to give less importance to the regions that have expression changes and, hence, have a less contribution to the final classification. The effect that expression changes have in different local areas is learned from a training set. The learned weights are then applied to the independent test face images. This approach is further extended in [7], where the weights are not learned but set inverse proportional to the difference in expression between the training and testing local regions.

More recently, this approach defined in this section has been extended to handle the problem of pose variations and to work with video sequences rather than simple stills [7]. Pose variations are again handled using Gaussian mixture models representing these variations. This algorithm has been shown to perform better than global approaches and voting strategies as well.

Another recent extension of this approach is given in [8], where the authors take advantage of the structure of the local areas by modeling it as a graph. This can be seen as a combination of the methods defined in the preceding section.

Extracting Sparse Components

Generally speaking, in the approaches defined above, there are infinitely many possible ways to divide a face image into a set of components. The question is *what is the optimal division?* The answer to this question will depend on the optimality criterion choose. When the separation is defined for a face recognition algorithm, the components are usually selected to include local regions that keep most of the main characteristics across the faces. This includes dividing the face into regions separating eyes, nose, and mouth or dividing

the face into equal local patches. On the other hand, if our goal is to represent (instead of discriminate) the faces *sparsely* using a component-based representation, one needs an algorithm that consider this alternate criterion. Such sparse representation of faces is required in many practical cases, since the complexity of representing several kinds of faces can simply be done by finding invariant component-based representations. An algorithm defined for this purpose is the nonnegative matrix factorization (NMF) [9].

In this approach, the parts of an object are learned automatically by the algorithm. The algorithm inputs the data (graylevel pixel values in the case of images) in matrix form, **V**, with each column representing an image. The goal is to factorize this matrix into basis vectors, **W**, and coefficients, **H**, such that none of the elements are negative, $\mathbf{V} = \mathbf{WH}$. To achive this, an Expectation Maximization (EM) like algorithm is proposed. Because of the nonnegativity constraints, the basis vectors become highly sparse leading to an efficient representation of the data matrix.

A major advantage of this algorithm is that it is able to extract the parts of the objects automatically according to their significance in the representation of the data. Since these parts are usually less variant under pose, illumination, and occlusions, one can design part-based recognition algorithms that are based on NMF features. An extension of this framework is given in [10].

Variety in Features

Some of the component-based algorithms defined in the literature, differ in the features that they use to represent the local regions. One of them is proposed in [11] where the authors extract Fourier Bessel coefficients of the local regions.

Three local regions around the eyes (left eye, between eyes, right eye) are cropped after automatically locating the face and eyes. The illumination changes across each local part are removed by means of an image normalization. If this corresponds to a region that has constant luminance, it is eliminated. This means that the occluded regions are eliminated. To extract the features, the local image patches are transformed to the polar frequency domain using the Fourier-Bessel Transform (FBT). This feature representation is used since the noise is eliminated using a subset of 372 FB coefficients. Using these coefficients the dissimilarities between each image with all the other images in the training set is calculated. Then, pseudo Fisher Linear Discriminant method (LDA) is used to classify the images [12]. The test results on FERET dataset [13] show that the proposed algorithm outperforms local approaches such as local polar Fourier Transform (PFT, which uses the Fourier transform instead of Fourier-Bessel), EBGM and global algorithms such as Principal Component Analysis (PCA), LDA, and global PFT. The results indicated that the proposed algorithm is sensitive to age and illumination changes more than expression changes.

The proposed algorithm is also tested with respect to its robustness to artificial occlusions, which is a drawback. Results on real occlusions are still needed. For this purpose 50% of the face was synthetically occluded with the graylevel information of those pixels equated to zero. In this experiment, the performance of local FBT was quite robust.

The authors further tested the significance of the localization error generated by their fully automatic system. They showed that the global approach is more affected by these errors than the local ones. And the local FBT performance was affected by up to 20% in the tests including age, expression, and illumination. This shows the importance of the localization errors in face recognition.

Other than changing the representation domain (pixel to polar frequency) of a local region, some algorithms use geometric features to represent the face components. One of these algorithms, which was already mentioned above, is the Face-ARG matching algorithm [8]. This algorithm uses the local information by extracting an Attributed Relational Graph (ARG). This graph is defined for each face image using a connected set of lines outlining the facial features. An important feature of the algorithm is that it uses only a single image for training, i.e., to extract the ARG. The testing is done using a partial match over the graph which was able to handle local changes and partial occlusions. The only disadvantage of the algorithm is the complexity of the matching. This is because the matching defined over the graph should also handle subset matching to deal with occlusions. The algorithm was able to obtain better recognition results on AR dataset [6] over most of the well-known algorithms such as nearest neighbor classification, PCA, and NMF.

Combined Face Detection and Identification

Not only face recognition but also face detection can be addressed using a component-based approach. In [14], the authors propose one such combined framework. They used a layered framework where the first layer is a component-based face detection module. This module consists of component classifiers specially trained for detecting facial components. Each detector outputs the most probable score and the x, y location of the corresponding component. A detection combination classifier receiving this data makes the final decision regarding the existence of a face in the image. If a face is detected, the obtained part-based regions are classified for identification in the second layer of the algorithm. Similar to the face detection module, the scores obtained from each part-based identification module are merged using a combination identity classifier, providing the final decision.

The training procedure in the face detection layer works as follows: First, 14 points on the face are selected as the center of the interest points. Then, starting from a rectangular region of predetermined size around each point, the interest region is increased until a minimum cross-validation error is obtained. On an independent testing set the trained face detector is able to outperform a detector that is exclusively based on global features.

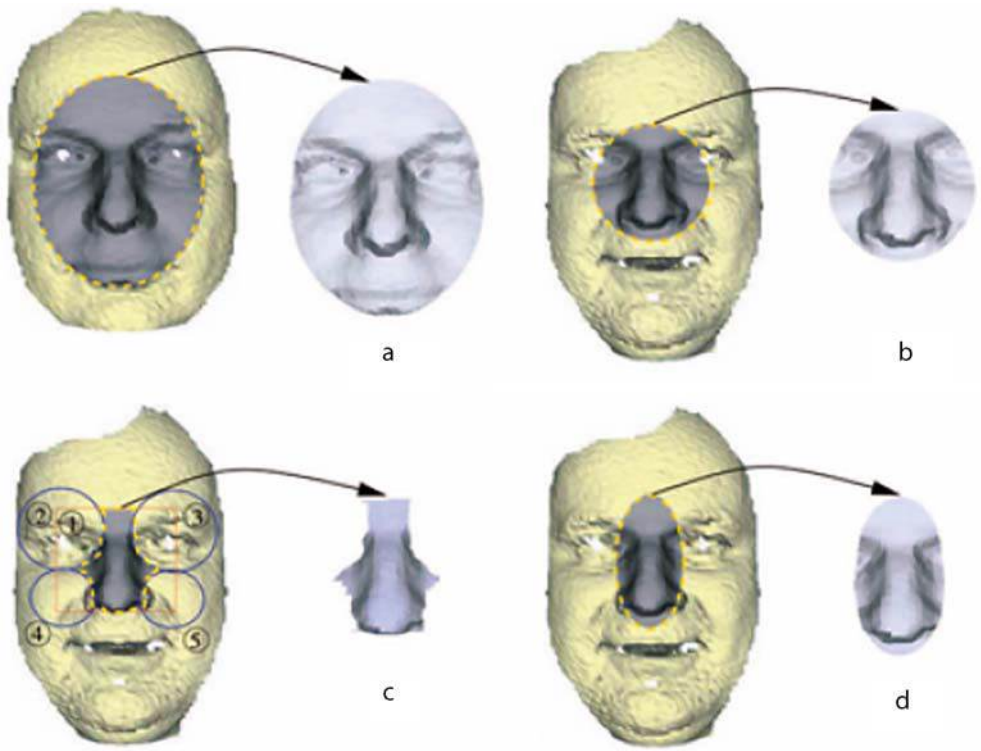
Once the size of the rectangular local region is determined in the detection layer, a linear Support Vector Machine (SVM) is trained for identification purposes. In this training procedure, 7040 synthetic face images are generated from a 3D model constructed for each individual using only three images, i.e., a frontal, a 45° rotated to the right, and a side face. Testing is held using 200 images from a total of 10 people. Images are recorded in different days and with different cameras. The proposed algorithm is compared to global face identification systems such as PCA, LDA, and a SVM with polynomial kernel. The authors further tested their combination face identifier using linear SVM with respect to possible combination such as majority vote, maximum product and maximum sum. All these combination scenarios perform better than the global methods with the linear SVM-based identifier combination. Specifically the linear SVM based on part-based region identification performed 89.25%, whereas the PCA, LDA and SVM based on global features obtained 61, 52, and 63%, respectively.

3D Face Recognition

The importance of part-based approaches in face recognition is also clear in algorithms defined to do recognition from 3D ▶ range scans [15]. In this paper, authors extend the part-based comparisons in 2D images to 3D range scans. The main idea in the paper is that the nose region is usually stable when the subjects have expression changes. Hence three different possible local regions are extracted from the 3D structure. Figure 4 shows these regions which correspond to a circular region centered around the nose, Fig. 4(b), a region exclusively dedicated to the nose (i.e., including the curvatures around the nose region), Fig. 4(c), and a larger, ellipsoidal region including the nose and its contextual information, Fig. 4(d). The experimental results obtained over expression variant faces show that using the 3D structure of the nose outperforms the use of the whole face. Several combination schemes are proposed to improve the results obtained by the local parts. Above all, the combination of the match scores obtained from the nose and ellipsoidal nose region performs the best for nonneutral expressions. This shows that the stability of the selected parts (in this case the nose region) can provide a moderate to large improvement in the performance in face recognition systems.

Object Detection

The part-based detection algorithms are also used in object detection in [16]. This approach depends on four major stages. In the first stage, the authors build a vocabulary of the parts to represent the images. This is done by applying the Forstner interest operator which detects intersection of lines and centers of circular patterns. Then, small image patches around the interest points are extracted. These image patches are clustered together to obtain a more compact set of image patches which is the part vocabulary. Second, each image is represented by a set of binary features stating whether the parts in the vocabulary are in the current image. The image representation also includes the spatial location of the parts in the image. This is done by calculating the relative distance and angular displacement between the parts and representing these



Face Recognition, Component-Based. **Figure 4** This figure shows three different nose regions extracted from the range scan shown in **(a)**. **(b)** corresponds to surface around the nose region, **(c)** is the region exclusively including the nose, and **(d)** is the surface of the ellipsoidal region around nose. © IEEE 2006.

in 5 and 4 bin histograms, respectively. In the third stage, the authors propose to learn a classifier by means of the sparse network of winnows learning architecture [17]. Since the feature representation is usually sparse this learning procedure is preferred. In the final stage, a classifier activation map is build to detect the objects in the test images. This is simply obtained by sliding the learned classifier with a window over the image and assign 0 whenever a negative activation occurs, and assigning the actual activation value otherwise. Once all the possible candidates have been identified over the whole image, neighborhood suppression is employed to eliminate false positives.

An extension to handle scale changes across the images is held using an image pyramid. At the end, the authors show the superiority of part-based approach with several tests over single scale or varying scale images.

SIFT Features

SIFT features are also applied to template matching problems because of the representational power of the component-based algorithms. In template matching the goal is to retrieve an object from a set of images. This problem needs to deal with local appearance variations, partial occlusions, and scale changes. To be robust to these variations a part-based approach can be used [18]. Due to the variety of images that a single object can generate, the training set that we need to learn is usually large. The complexity of the algorithms not only increase with the number of samples in the training set but also with the slow template matching algorithms, such as calculating the sum of square distance (SSD) or the normalized cross-correlation (NCC) between images [12]. To eliminate this computational complexity [18] proposed to use rectangular filters that are usually employed for fast image filtering in integral image representations. Furthermore, the complexity of the training set is reduced by means of a part-based representation instead of global templates.

In this approach each image is divided into a set of patches, where each of them is filtered with a set of rectangle filters. This process usually leads to a smaller number of feature representation. A subset of these

features are then selected using a saliency threshold specified by the user. This is done such that the features that are closer to zero are eliminated, since these are mostly related to patches that have constant pixel and thus carry little classification information. Furthermore, a weighted approach is used to decrease the significance of the patches that have more appearance changes across the images, i.e., the mouth region in the face. To handle partial occlusions the most similar parts are used during the matching process. The variable scale of the images is given by the property of the rectangular filters. They have shown that this method can robustly and accurately classify faces very fast under partial occlusions, variable expression, and different scales.

Similar problems to those observed in template matching are also seen in image retrieval applications. Therefore, most of the algorithms defined for image retrieval extract local features from the images [19]. Of late, one of the most popular techniques used in this process are the Shift Invariant Feature Transform (SIFT) features. In this algorithm, each image is represented by the set of directional gradient vectors on local regions. Using a similar idea, [20] proposed PCA-SIFT which is shown to be a compact, fast, and accurate representation for faces. The key idea is to use PCA to describe the gradient information located around the keypoints selected by SIFT. Using PCA, the authors show that a small number of basis images (20) are enough to represent these gradient based images. The Receiver Operating Characteristic (ROC) curves would be inappropriate to analyze this algorithm, since this corresponds to a detection problem with a large number of false positives rather than to a classification one. Thus, recall versus 1-precision curves which compared *correct positives/number of positives* (recall) with *false positives/number of matches* (1- precision) (normalized measurement) are employed. Using these analysis, the authors show that PCA-SIFT performs better than SIFT in several different scenarios such as, added noise, rotation and scale changes, projective warp, and reduced brightness. The conclusion from this algorithm is that SIFT features may include noise in the description of the local information, whereas using PCA on the local graylevel eliminates these, facilitating the final classification of faces.

A Comparison of Different Approaches

A recent comparative study for the local matching approach in face recognition is presented in [21]. In this review, methods are categorized according to alignment/partitioning algorithms, local feature representations, and classification combinations. Alignment/partitioning algorithms are also divided within themselves into three subclasses. The first of them uses the local components defining a face, such as eyes, nose, and mouth. It is argued that these features may not be appropriate when one wants to consider the relationship between components. Another set of algorithms uses ▶ Face Warping in order to obtain shape-free graylevel information. Instead of removing the shape, the third set of algorithms eliminate the affine transformation between the images and then employs a part-based representation.

The authors discuss several local feature representation types. These include the Eigenfeatures, Gabor features, and local binary pattern features. Eigenfeatures use the pixel level information and eliminate the noise and represent the images using an orthonormal bases. Gabor features are extracted using a set of Gabor jets over the parts and represent each component according to the output obtained with filtering. Local binary patterns are obtained by calculating the binary patterns around an interest point for a given radius.

Furthermore, the algorithms are cataloged according to the classification method and the combination of the local cues. Local features are either simply concatenated into a global feature or combined with weights. An alternative method is to use a weighted combination of the classifiers defined by the local parts. While in other cases, the sum rule or Borda count (i.e., the classifier that has the largest votes) may be used to combine the classification decisions.

The experiments are carried out with the FERET and the AR face databases and show that LBP performs the best when the images are used with no illumination change, while the Gabor jets are better when lighting is considered. Another experiment is conducted to learn the best components for face recognition, revealing that the nose region generally outperforms the others. This may be due to the fact that the nose region is the most stable part under changing expression.

The authors also consider the difference between local region approaches (i.e., regions around the fiducial points) and the use of local components (i.e., components that are the parts of the face image divided equally, similar to a grid structure). They conclude that the local region approach may perform

better since the shape information is also used implicitly in the process.

Another question is whether or not to use other regions of the face, such as the cheeks and the forehead. In [21], it is shown that although the use of such regions may degrade the performance of the global approaches such as eigenfeatures, and their use in local methods, as in LBP and Gabor jet, generally improves the final recognition rate.

Based on these experiments and observations, [21] defines a new component-based approach that uses Gabor jets to extract features from local regions at several scales and frequency values. They combine the classification results (where the similarity is obtained by normalized inner products) with the Borda count. This approach was able to outperform the others in a test using the FERET database.

Summary

In this chapter, the authors have reviewed some of most known and recent works on component-based face recognition. All of the algorithms summarized above are defined to handle one or more of the problems of face recognition, as for example, imprecise localization, pose, illumination, expression, and occlusion.

To do that some algorithms such as DLA and EBGM are defined to be invariant to localization of the faces. Alternatively, we can model the image variations or use a sparse representation. NMF tries to extract a sparse local representation for the components of the dataset. Some approaches such as FBT and Face-ARG use different features to represent the local regions. The expression varying 3D scans showed that using the nose as the local component improves the recognition from 3D data. While another algorithm defined to do recognition from a single image learns all possible image changes when possible and uses weights to determine which regions are most robust to variations elsewhere. These changes can be modeled using Gaussian distributions, e.g., a mixture of Gaussians representing the variations when the face is imprecisely localized or when an expression varies the brightness of the pixels in a patch. A similar approach is also defined for face recognition from video where the pose changes are also considered. Some other algorithms use the local information both for face detection and identification, whereas alternatives are defined for generic object detection. Finally, the authors have summarized the results of a recent comparison.

All these algorithms have one common thread: considering the images as a combined set of components. This is mainly because of the stability of the local components over possible image variations. The use of component-based algorithms is to date one of the most used approaches in classification and identification of 2D and 3D faces.

Related Entries

- ▶ Face Alignment
- ▶ Face Localization
- ▶ Face Pose Analysis
- ▶ Face Recognition, 3D-Based

References

1. Martinez, A.M.: Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. IEEE Trans. Pattern Anal. Mach. Intell. **24**(6), 748–763 (2002)
2. P. Belhumeur, D.K.: What is the set of images of an object under all possible illumination conditions? Int. J. Comput. Vis. **28**(3), 245–260 (1998)
3. Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., Vandermalsburg, C., Wurtz, R.P., Konen, W.: Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput. **42**(3), 300–311 (1993)
4. Wiskott, L., Fellous, J.M., Kruger, N., vonderMalsburg, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 775–779 (1997)
5. Martinez, A.M.: Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head, SC, USA (2000)
6. Martinez, A., Benavente, R.: The AR-face database. Tech. rep., CVC Tech. Report # 24 (1998)
7. Zhang, Y.B., Martinez, A.M.: A weighted probabilistic approach to face recognition from multiple images and video sequences. Image Vis. Comput. **24**(6), 626–638 (2006)
8. Park, B.G., Lee, K.M., Lee, S.U.: Face recognition using Face-ARG matching. IEEE Trans. Pattern Anal. Mach. Intell. **27**(12), 1982–1988 (2005)
9. Lee, D.D., Seung, H.S.: Learning the parts of objects by non-negative matrix factorization. Nature **401**(6755), 788–791 (1999)
10. Guillamet, D., Bressan, M., Vitrià, J.: Weighted non-negative matrix factorization for local representations. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, pp. 942–947 (2001)
11. Zana, Y., Cesar, R.M., Feris, R., Turk, M.: Local approach for face verification in polar frequency domain. Image Vis. Comput. **24**(8), 904–913 (2006)

12. Fukunaga, K.: Introduction to Statistical Pattern Recognition. Academic Press, New York (1990)
13. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The feret database and evaluation procedure for face recognition algorithms. Image Vis. Comput. **16**(5), 295–306 (1998)
14. Heisele, B., Serre, T., Poggio, T.: A component-based framework for face detection and identification. Int. J. Comput. Vis. **74**(2), 167–181 (2007)
15. Chang, K.I., Bowyer, K.W., Flynn, P.J.: Multiple nose region matching for 3D face recognition under varying facial expression. IEEE Trans. Pattern Anal. Mach. Intell. **28**(10), 1695–1700 (2006)
16. Agarwal, S., Awan, A., Roth, D.: Learning to detect objects in images via a sparse, part-based representation. IEEE Trans. Pattern Anal. Mach. Intell. **26**(11), 1475–1490 (2004)
17. Carlson, A.J., Cumby, C., Rosen, J., Roth, D.: The snow learning architecture. Tech. rep., Technical Report UIUCDCS-R-99-2101, Computer Science Department, University of Illinois at Urbana-Champaign (1999)
18. Guo, G., Dyer, C.: Patch-based image correlation with rapid filtering. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Minneapolis, Minnesola, USA (2007)
19. Martinez, A.M.: Face image retrieval using hmms. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (Workshop) Ft. Collins, CO, USA (1999)
20. Ke, Y., Sukthankar, R.: Pca-sift: A more distinctive representation for local image descriptors. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC, USA (2004)
21. Zou, J., Ji, Q., Nagy, G.: A comparative study of local matching approach for face recognition. IEEE Trans. Image Process. **16**(10), 2617–2628 (2007)

Face Recognition, Geometric vs. Appearance-Based

Lior Wolf
The Blavatnik School of Computer Science, Tel-Aviv University, Israel

Synonyms

Features vs. Templates; Shape vs. Texture

Definition

In 2D face recognition, images are often represented either by their geometric structure, or by encoding their intensity values. A geometric representation is obtained by transforming the image into geometric primitives such as points and curves. This is done, for example, by locating distinctive features such as eyes, mouth, nose, and chin, and measuring their relative position, width, and possibly other parameters. Appearance-based representation is based on recording various statistics of the pixels' values within the face image. Examples include: recording the intensities of the image as 2D arrays called templates and computing histograms of edge detectors' outputs.

Introduction

Face identification systems are challenged by variations in head pose, camera viewpoint, image resolution, illumination, and facial expression, as well as by longer-term changes to the hair, skin, and head's structure. The geometric approach, which transforms a face image into simple geometric primitives, holds the promise of being invariant to illumination and almost invariant to time-induced changes. Using well-understood ▶ Multiple View Geometry techniques, it can also be made practically invariant to minor pose differences, camera viewpoint changes, and image resolution. In addition, the geometric approach has the advantage that a geometric match is easy to interpret.

Inspite of their intuitive and seemingly precise nature, geometric face recognition techniques have been largely replaced by appearance-based techniques. In these techniques, image representations, which are directly computed from the pixel-intensities are compared to estimate similarities between images, or fed into ▶ classifiers that determine the identity of the person in the image.

Even though the appearance-based techniques are cleverly designed and engineered, they lack the rigorous nature of the geometric approach. When an appearance-based classifier determines a false identify or wrongly detects a match between two persons, it is often hard to understand why this happens. Nevertheless, in 1993 Brunelli and Poggio [1] have shown that a generic appearance-based method outperforms a simple geometric-based method on the same dataset, and contradicting evidence to their finding has been scarce.

Shape-based methods

The pioneering work of Kanade [2], which was among the first modern approaches for automatic identification of face images, used the geometric approach. His system, similarly to following contributions, starts by identifying the locations of dominant facial features such as the eyes' corners, the nostrils' center, and the mouth's extremities. This detection step is the most challenging step, and the detected features are selected to be both discriminative and easily detectable. Other desired properties include invariance to lighting conditions and to facial expression.

The second step consists of defining a vector of measurements that are used as a face signature. Kanade [2] uses 16 such measurements, which include ratios of distances (e.g., the ratio of the distance between the eyes and the width of the face), various facial angles, the chin's curvature and more. Brunelli and Poggio [1] use a different set of 35 features, which include eyebrow thickness, shape and position, nose and mouth position, facial width at several heights and the shape of the chin (Fig. 1).

The third step consists of comparing two faces, or more generally, learning a classifier that can identify the various face classes. Kanade uses a simple Euclidean distance to compare two signatures. The later work [1] employs principle component analysis (PCA) of various dimensionality. Peak performance is obtained with the maximal dimensionality, where the distance between signatures becomes their Euclidean distance.

An improvement to this basic three-step face recognition scheme can be obtained by taking advantage of the increase in accuracy of facial feature detection algorithms (e.g., [3]) by designing or learning more discriminative signatures, and by introducing modern machine-learning techniques to learn distance functions between signatures or to train better classifiers. To our knowledge, little work has been done to demonstrate any of these improvements.

An alternative framework [4] for extracting geometric primitives from face images is to ignore the high-level structure of faces (i.e., as composed of identifiable eyes, nose, etc.) and instead of transforming the image into line drawings containing many line segments. First, edge detectors are applied to detect edge pixels, followed by a morphological thinning operator. Then, a line fitting process [5] is used to break continuous edge curves into several short line segments. The set of obtained segments (each represented by its endpoints) constitutes the signature of the face image.

Face Recognition, Geometric vs. Appearance-Based.
Figure 1 Some of the geometrical features used in [1], including eyebrow thickness and vertical position at the eye center position, nose position and width, the mouth's vertical position and width, height of lips, eleven radii describing the chin's shape, width of face at nose height, and its width halfway between the nose and the eyes. Other features which are not shown include a description of the left eyebrow's shape. Figure adapted from Fig. 6 of [1].

In order to compare two such signatures, a distance measure between two sets of line segments is required. In [4] an elaborate such measure is proposed which considers the fact that two similar line segments can differ in length, may be tilted or be parallel, and that some matching line segments may be missing.

Appearance-based methods

Even though the terms shape and geometry are often used synonymously, we should not be misled to

assume that appearance-based methods do not encode the face's shape. Indeed, the location of the facial features and their shape contribute much to the variation in appearance between persons. The other identity-based source of appearance variation between persons is the facial texture, which includes elements that are typically not encoded by the geometry-based methods such as skin tone, facial hair, freckles, scars, and wrinkles.

Most appearance-based techniques share similar stages or components, which are sometimes intertwined:

Normalization	During this initial stage, images may be scaled such that the area of the detected face is approximately constant. In addition, faces can be warped to a fixed reference image, see Fig. 2(b). A cropping mask is sometimes applied to remove the image boundary region, which typically includes hair and background elements. Furthermore, histogram equalization or other means of reducing the effects of illumination may be applied.;
Signature generation	The normalized face image is being processed according to the specific algorithm at hand and a signature is created. For example, a local texture descriptor may be computed at each pixel, and the histogram of this descriptor can be used as a signature.;
Learning or Classification	A classifier is trained to distinguish between the various persons in the database, or a distance function is learned to estimate the likelihood of two signatures belonging to the same person. These are used to classify the new images not used during training.

The most basic signatures are based on templates derived directly from the image. Variations may include using image derivatives instead of image intensities, or otherwise normalizing the intensities to reduce the effect of illumination. Another class of variations consists of using several templates (components) out of each face image [1, 6], and combining the matching score of each component in the final score.

Much work has been put into defining discriminative face signatures based on local texture descriptors. Local binary patterns (LBP) have shown to be extremely effective for face recognition [7]. The most simple form of LBP is created at a particular pixel location by threshholding the 3 \mathop{\times} 3 neighborhood surrounding the pixel with the central pixel's intensity value, and treating the subsequent pattern of 8 bits as a binary number (Fig. 3). A histogram of these binary numbers in a predefined region is then used to encode the appearance of that region. Typically, a distinction is made between uniform binary patterns, which are those binary patterns that have at most 2 transition from 0 to 1, and the rest of the patterns. For example, 1000111 is a uniform binary pattern while 1001010 is not. The frequency of all uniform LBPs is estimated, while all nonuniform LBPs, which are typically around 10% of patterns in an image, are treated as equivalent and given only one histogram bin. LBP representation for a given face image is generated by dividing the image into a grid of windows and computing histograms of the LBPs within each window. The concatenation of all these histograms constitutes the signature of the image.

A large body of literature exists on the proper way of learning classifiers and distances for face recognition. The PCA-based "eigenfaces" method [8] and the LDA based "fisherfaces" method [9] have been the first in a constant stream of work. It is the author's experience that for modern descriptors such as LBP, All-Pairs Support Vector Machine [10] performs well in the task of biometric identification.

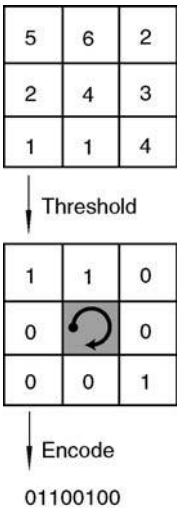
Recently, some effort has been devoted to the estimation of visual similarities between two unseen images, and such methods have been applied to determine whether two images belong to the same person. One method [11] that has shown good results for uncontrolled imaging conditions uses Randomized Decision Trees [12] and Support Vector Machines. In the first image of the pair, image patches (fragments of the image) are selected at random locations. For each patch the most similar patch in the second image is searched at a nearby image location. A decision tree is trained to distinguish between pairs arising from matching images and those arising from nonmatching images. Given a pair of unseen images, a Support Vector Machine classifier is used to determine if they match by aggregating the Decision Tree output of many image patches.

Face Recognition, Geometric vs. Appearance-Based. **Figure 2** Various methods of alignment. (**a**) The original image. (**b**) The congealing method [18] is used to align the image to a semi-frontal view. (**c**) Facial feature points located by the system of [19]. The detected feature points can be used to create a geometric signature or to align the image as a preprocessing step to an appearance-based approach.

Hybrid methods

Some face recognition methods combine appearance and geometry. In [13], the authors employ a face tracking method called Active Appearance Model to locate a set of feature points in and around the face. An example of similar features captured by a subsequent system, can be seen in Fig. 2(c). The located feature points are used to create three signatures that are combined during the recognition process. The first signature encodes the locations of the detected feature points; The second signature encodes the gray values of the image after it is ▶ warped such that the feature points are mapped to the mean location of those points in the training datasets. The third signature encodes the local appearance around the detected feature points.

The Elastic Bunch Graph Matching system [14] uses local Gabor-wavelet based detectors that are connected through a simple spring model to locate facial features in a new image. The detected fiducial points are used to position a finer grid of points on the face image, and the response of various Gabor wavelets on the grid points is recorded to describe their local appearance. The matching score between two facial grids takes into account both the locations of the grid points and their appearance. This method performs well (for its time), however, the matching process is slow.



Face Recognition, Geometric vs. Appearance-Based.
Figure 3 The LBP image-texture descriptor is computed locally at each pixel location. It considers a small neighborhood of a pixel, and thresholds all values by the central pixel's value. The bits which represent the comparison results are then transformed into a binary number. The histogram of these numbers (the vector containing the frequency of each binary number in the image) is used as a signature describing the texture of the image.

A much faster hybrid method [15] uses coupled Gaussian mixture models to locate eyes, nose tip, and mouth in images of varying pose. Five SIFT appearance descriptors [16] are computed in regions around the detected features and in between the eyes.

Summary

Appearance-based methods currently dominant the general field of object recognition, where more classical methods based on analysis of relative positions of corners and other feature points have been mostly abandoned. Furthermore, there is evidence that the same image descriptors can be used for both object recognition and face identification [7, 15, 17]. It is therefore not surprising that the leading face recognition methods are also appearance based.

However, human faces differ from most objects studied in object recognition in that they have a well defined structure. It is possible that the major disadvantage of geometric face recognition is the lack of robust facial feature detectors. The advent of new detection techniques may reignite the interest in those methods.

Related Entries

- ▶ Face Recognition, Component-based
- ▶ Face Recognition, Overview
- ▶ Face Recognition, Sketch-based
- ▶ Face Tracking

References

1. Brunelli, R., Poggio, T.: Face recognition: Features versus templates. IEEE Trans. Pattern Anal. Mach. Intell. **15**(10), 1042–1052 (1993)
2. Kanade, T.: Picture processing system by computer complex and recognition of human faces. Ph.D. thesis, Kyoto University (1973)
3. Ding, L., Martinez, A.: Precise detailed detection of faces and facial features. In: Proceedings of IEEE Computer Vision and Pattern Recognition, pp. 1–7 (2008)
4. Gao, Y., Leung, M.: Face recognition using line edge map. IEEE Trans. Pattern Anal. Mach. Intell. **24**(6), 764–779 (2002)
5. Leung, M.K., Yang, Y.H.: Dynamic two-strip algorithm in curve fitting. Pattern Recognit. **23**(1–2), 69–79 (1990)
6. Heisele, B., Serre, T., Poggio, T.: A component-based framework for face detection and identification. Int. J. Comput. Vis. **74**(2), 167–181 (2007)
7. Ahonen, T., Hadid, A., Pietikainen, M.: Face Description with Local Binary Patterns: Application to Face Recognition. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 2037–2041 (2006)
8. Turk, M., Pentland, A.: Eigenfaces for Recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
9. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)
10. Allwein, E.L., Schapire, R.E., Singer, Y.: Reducing multiclass to binary: a unifying approach for margin classifiers. J. Mach. Learn. Res. **1**, 113–141 (2001)
11. Nowak, E., Jurie, F.: Learning visual similarity measures for comparing never seen objects. In: IEEE Conference on Computer Vision and Pattern Recognition (2007)
12. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. J Mach. Learn. Res. **36**(1), 3–42 (2006)
13. Lanitis, A., Taylor, C.J., Cootes, T.F.: A unified approach to coding and interpreting face images. In: Proceedings of the International Conference on Computer Vision, pp. 368–374. IEEE Computer Society, Washington, DC, USA (1995)
14. Wiskott, L., Fellous, J.M., Kröger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 775–779 (1997)

15. Sivic, J., Everingham, M., Zisserman, A.: Person spotting: Video shot retrieval for face sets. In: 4th International Conference on Image and Video Retrieval, pp. 226–236 (2005)
16. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
17. Meyers, E., Wolf, L.: Using biologically inspired features for face processing. Int. J. Comput. Vis. **76**(1), 93–104 (2008)
18. Huang, G., Jain, V., Learned-Miller, E.: Unsupervised joint alignment of complex images. Computer Vision, In: IEEE International Conference pp. 1–8 (2007)
19. Zhou, Y., Gu, L., Zhang, H.J.: Bayesian tangent shape model: estimating shape and pose parameters via bayesian inference. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. I–109–I–116 (2003)

Face Recognition, Near-Infrared

Stan Z. Li, Dong Yi
Biometrics and Security Research & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China

Synonyms

Near-infrared image based face recognition; Face recognition in near-infrared spectrum

Definition

Near-infrared (NIR) based face recognition, as opposed to the conventional *visible light* (VIS) based, is an effective approach n overcoming the impact of illumination changes on face recognition. It uses a special purpose imaging capture hardware, in which active NIR lights mounted around the camera lens illuminate the face from near frontal direction and an NIR camera captures front-lighted NIR face images. This is similar to a camera flash but the imaging is done in the *invisible* NIR spectrum. With such NIR face images, problems caused by uncertainties in uncontrollable environmental ▶ illumination are minimized, and difficulties in building the face matching engine is much alleviated. The NIR approach usually achieves significantly higher performance than the VIS approach.

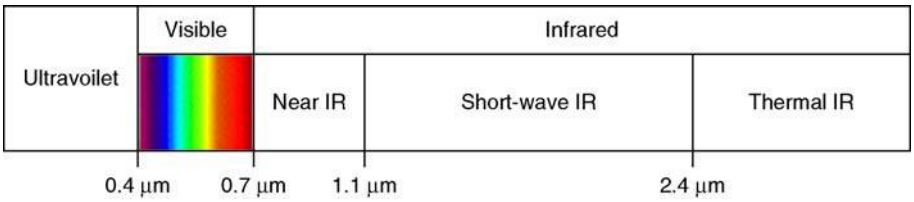
Introduction

Face recognition should be performed based on intrinsic factors of the face, related to the 3D shape and albedo of the facial surface. In contrast, extrinsic factors, including eyeglasses, hairstyle, expression, posture, and lighting should be minimized because they make distributions of face data highly complex.

Among the aforementioned extrinsic factors, problems with uncontrolled environmental (ambient) illumination is the important issue [1]. Illumination direction is the most critical of all [2]. From the end-user point of view, a biometric system should adapt to the environment. However, face recognition systems based on face images captured in visible light (VIS) spectrum are compromised of changes in environmental illumination, even for cooperative user applications with frontal faces captured indoor. Numerous publications exist for modeling and normalizing face illumination conditions. They are found to improve recognition performance, but have not led to a face recognition method which is illumination independent.

3D face recognition provides a solution to the illumination problem. Disadvantages of it include increased cost, lesser speed, and specular reflections. It is reported that the 3D methods do not necessarily produce better recognition results than the 2D methods [3].

Imaging and vision beyond the visible spectrum has recently received much attention in the computer vision community (e.g., [4]). Radiation spectrum ranges are shown in Fig. 1. Instead of ultraviolet radiation which is harmful to the human body, thermal-infrared and near infrared (NIR) imagery are employed for face recognition applications. Such "invisible" spectrum imaging technologies are effective in dealing with uncontrolled illumination. This is because they work in different bands, from the conventional VIS imaging to many visual effects, as encountered in conventional illumination changes can be eliminated. Disadvantages of the FIR approach include instability due to environmental temperature, emotional and health conditions, and poor eye localization accuracy [5]. The use of active near infrared (NIR) imaging brings



Face Recognition, Near-Infrared. **Figure 1** Radiation spectrum ranges.

a new dimension for face detection and recognition [6, 7, 8, 9, 10, 11].

The key part in the NIR face recognition approach is a special purpose image capture hardware system [10, 11]. It uses active NIR illuminators, for example, ▶ light-emitting diodes (LEDs), mounted around the camera lens to illuminate the face from near front direction and then capture front-lighted NIR face images. This is similar to a camera flash but as the NIR lighting works in *invisible* NIR spectrum it is non-intrusive to human eyes.

An NIR face image with frontal illumination is subject mainly to an approximately monotonic transform in the gray tone, and problems caused by uncertain environmental illumination are minimized. Therefore, the face detection and matching algorithms need to cope with this degree of illumination changes mainly. This is much less difficult than the problems with conventional VIS face images.

The NIR approach usually achieves significant higher performance than the VIS approach [11]. The use of NIR techniques leads to highly accurate and fast face recognition systems for cooperative face recognition applications, indoors [10, 11] and outdoors [12]. The use of NIR face images for biometrics is now being evaluated by NIST [13].

A limitation, however, is that both enrollment and the query face images should be of the NIR type, which similar to the requirement for 3D face recognition. Methods for matching the NIR query and VIS target images that are required for photo IDs, are being developed [14].

NIR Imaging Hardware

The goal of making this special-purpose hardware is to overcome the problem arising from uncontrolled environmental light so as to produce face images of a good illumination condition for face recognition. "A good illumination condition" means that the lighting is from the frontal direction and the image has suitable pixel intensities.

This could be achieved by the following methods: (1) Active NIR light can be mounted (e.g., 850nm LEDs) around the camera lens to provide strong frontal lighting enough to override environmental light, and set a low camera exposure to produce a clear frontal-lighted face image. (2) a ▶ long-pass optical filter can be used to further minimize the remaining environmental lighting by cutting off visible light of wavelength shorter than 750nm. Fig. 2 illustrates an example of the hardware device, and resulting face images. The face in the images are illuminated by NIR LED light from the front and a lamp aside, in addition to other environmental light.

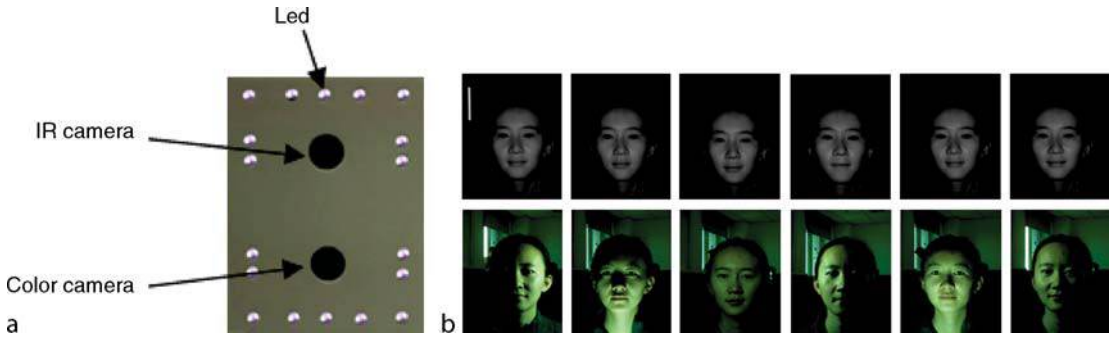
In outdoor environments, the sunlight contains much stronger NIR component than NIR LEDs. The hardware must be further designed to minimize influence of the sunlight to maintain the effect of the active NIR illumination. It could be enhanced by using a strong active NIR pulse illuminator and NIR camera, co-working in a synchronized manner [12].

Illumination Invariant Face Representation

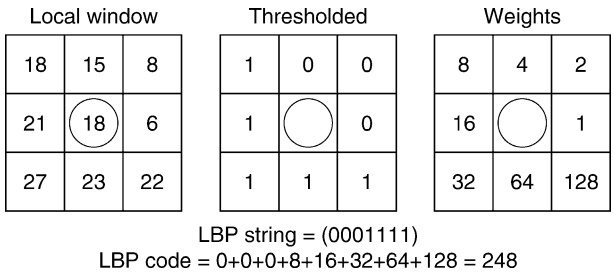
According to the ▶ Lambertian law, an image $I(x, y)$ under a point light source is formed according to the following equation

$$I(x, y) = \rho(x, y)\mathbf{n}(x, y)\mathbf{s} \quad (1)$$

where $\rho(x, y)$ is the albedo of the facial surface material at the point (x, y), $\mathbf{n}=(n_x, n_y, n_z)$ is the surface normal (a unit row vector) in the 3D space, and $\mathbf{s}=(s_x, s_y, s_z)$ is



Face Recognition, Near-Infrared. **Figure 2** An experimental active NIR imaging device (with an additional color camera), and NIR versus color images captured under different environmental lightings. While unfavorable lighting changes are obvious in the color images, they are almost unseen in the NIR images.



Face Recognition, Near-Infrared. **Figure 3** LBP code for 3x3 window.

the lighting direction (a column vector, with magnitude). Here, albedo $\rho(x, y)$ reflects the photometric properties of facial skin and hairs, and $\mathbf{n}(x, y)$ is the geometric shape of the face.

The LEDs mounted around the camera lens are approximately co-axial to the camera direction, and thus provide the best possible straight frontal lighting. In this case, the image can be approximated by

$$I(x, y) = \kappa\rho(x, y)n_z(x, y) \qquad (2)$$

where $n_z(x, y)$ is the depth information (2.5 map) that can be acquired by a range imaging system, and κ can be modeled as being monotonic to the distance between the face and the active light.

The degree of freedom due to the monotonic transform of κ may be compensated by applying some operator, such as local binary pattern (LBP), on the NIR image to produce a genuine illumination invariant face representation [11]. The basic form of the LBP operator is illustrated in Fig. 3. The binary bits describing a local 3 x 3 subwindow are generated by thresholding the 8 pixels in the surrounding locations by the gray value of its center; the feature vector is formed by concatenating the thresholded binary bits anti-clockwise. The LBP code does not change with any monotonic transform of the image. Therefore, applying an LBP operator to an active NIR image generates illumination invariant features for faces. A highly accurate face recognition system can then be built.

Summary

The NIR approach uses an active NIR imaging hardware to acquire front-illuminated face images, to overcome the problem of illumination variation that every face recognition system has to deal with. NIR face images have good properties and render extraction of illumination invariant face features for building accurate face recognition systems. The use of NIR face images for face biometrics is now being evaluated by NIST [13].

A limitation of the NIR approach, however, is that both enrollment and the query face images should be of the NIR type. Methods for matching the NIR query

and VIS target images that are required for photo IDs are yet to be developed.

Related Entries

- ▶ Face Recognition Overview
- ▶ Hyperspectral and Multispectral Biometrics
- ▶ Local Binary Pattern (LBP)

References

1. NIST: Face Recognition Vendor Tests (FRVT). http://www.frvt.org (2006)
2. Adini, Y., Moses, Y., Ullman, S.: Face recognition: The problem of compensating for changes in illumination direction. IEEE Trans. Pattern Anal. Mach. Intell. **19**, 721–732 (1997)
3. Chang, K.I., Bowyer, K.W., Flynn, P.J.: An evaluation of multimodal 2D+3D face biometrics. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 619–624 (2005)
4. OTCBVS. In: IEEE International Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum. (2004–2005)
5. Chen, X., Flynn, P.J., Bowyer, K.W.: Infra-red and visible-light face recognition. Comput. Vis. Image Understand. **99**, 332–358 (2005)
6. Dowdall, J., Pavlidis, I., Bebis, G.: Face detection in the near-IR spectrum. Image. Vis. Comput. **21**, 565–578 (2003)
7. Li, D.Y., Liao, W.H.: Facial feature detection in near-infrared images. In: Proceedings of fifthth International Conference on Computer Vision, Pattern Recognition and Image Processing, Cary, NC, pp. 26–30 (2003)
8. Pan, Z.H., Healey, G., Prasad, M., Tromberg, B.: Face recognition in hyperspectral images. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 1552–1560 (2003)
9. AuthenMetric Co. Ltd.: A Method for Face Image Acquisition and a Method and System for Face Recognition. Patent No.PCT/CN2004/000482 (2004)
10. Li, S.Z., His Face Team: AuthenMetric F1: A Highly Accurate and Fast Face Recognition System. ICCV2005 - Demos (2005)
11. Li, S.Z., Chu, R., Liao, S., Zhang., L.: Illumination invariant face recognition using near-infrared images. IEEE Trans. Pattern Anal. Mach. Intell. **26** (2007)
12. Yi, D., Liu, R., Chu, R., Liu, D., Wang, R., Li, S.Z.: "Outdoor face recognition using enhanced near infrared imaging". In: Proceedings of IAPR International Conference on Biometric, Seoul, Korea (2007)
13. NIST: Multiple Biometric Grand Challenge (MBGC). http://face.nist.gov/mbgc (2008)
14. Yi, D., Liu, R., Chu, R., Lei, Z., Li, S.Z.: Face matching between near infrared and visible light images. In: Proceedings of IAPR International Conference on Biometric, Seoul, Korea (2007)

Face Recognition, Overview

Aleix M. Martinez
Department of Electrical and Computer Engineering, Ohio State University, Columbus, OH, USA

Synonyms

Face Biometric; Face Identification; Face Verification

Definition

Face recognition is the science which involves the understanding of how the faces are recognized by biological systems and how this can be emulated by computer systems. Biological systems employ different types of visual sensors (i.e., eyes), which have been designed by nature to suit a certain environment where the agent lives. Similarly, computer systems employ different visual devices to capture and process faces as best indicated by each particular application. These sensors can be video cameras (e.g., a camcorder), infrared cameras, or among others, 3D scans. The essay reviews some of the most advanced computational approaches for face recognition defined till date.

Introduction

Many types of biometrics exist for identifying a person or verifying that a given individual is what he or she claims to be. Some of the biometrics result in quite reliable recognition and verification, but most are either intrusive to the individual or expensive (e.g., DNA or iris). Furthermore, many of the biometrics have raised reasonable questions about an individual's rights and personal freedom [1]. The systems that are typically considered less intrusive by people, are those based on the recognition of faces.

We are so used to seeing and recognizing faces that most people think computers should have such a capacity too. Computer face recognition allows devices to recognize and interact with users, allowing them to go beyond the boring and slow use of the keyboard and mouse. The face carries so much information that

people find it difficult to interact on the phone for long, forcing companies to include cameras and even video on cell phones – even if the most recorded sequence is perhaps a simple "Hi." Yet, people feel that the smile associated with a simple greeting is essential to start a good conversation or for carrying the real and intended thought of the messenger. The face also provides the identity of the speaker, avoiding the awkwardness of trying to identify someone by voice alone. This effect is now clearer than ever, with video chats becoming increasingly popular as high speed Internet becoming available at low costs to the general public.

It is thus not surprising that people still remain open to the possibility of having face recognition systems at home, work, or other places like at the ATM. One concern with this technology is to make sure that the biometrics of the individuals cannot be stolen. Imagine a scenario where a hacker steals information from a database of faces and then employs this to hack other computer, systems or institutions with a stolen identity. A password or an ID card can be changed, but a face cannot be. To address these concerns, researchers in face recognition are developing mechanisms to encrypt personal biometrics. One classical solution is to define a mapping function which maps a face image into a single instance (e.g., feature vector). The trick is to use a function whose inverse mapping is not unique (i.e., the inverse mapping results in multiple solutions) unless you know the encryption key [2]. In face recognition, we may even be able to eliminate the need for the encryption key. This can be achieved by defining a recognition algorithm in the encrypted space. This is possible because of the uniqueness of direct mapping. This means we can perform face recognition even when the understanding the information stored in our own database is not possible (i.e., in the sense, the image do not have meaning for the human visual system any longer). This could mean that general databases of faces could be shared by several institutions, because these cannot abuse its contents. Also, if the database of face images is stolen, unauthorized users would not be able to make sense of its data. These security protocols generally make face recognition systems more acceptable by the general population.

Perhaps the most important disadvantage of face recognition is that it cannot provide as accurate an identification as other biometrics, and definitely not as accurate as DNA or iris. Nonetheless, in a large number of applications such a secure analysis is not needed. One of the most classical examples is in human–computer interaction – the cell phone example given above being but one example of its potential uses. Another typical example application is wherever individuals need to gain access to restricted areas within a company. This is of particular use where the employees are known a priori. One well-known case is in airports, where not all personnel have access to the runway or planes. A related application is for costumer verification, for example for airline tickets, where the manual picture to face check is known to be flawed. In the 2008 summer Olympics, the organizer gave the opportunity to attendees of the inaugural ceremony to attach a picture of their face to their tickets. At the ceremony, the holders of these tickets were asked to look at a camera and a computer compared the face of the ticket holder with that of the buyer. A mismatch prompts the organizers to request additional information to demonstrate that the identity of the ticket holder and buyer is the same.

What makes all these applications and many others possible is the tremendous advances that have been accomplished in the past years in the area of computer algorithms for automatic face recognition. Current systems are able to recognize faces under a large number of variations; sometimes overpassing human performance. However, to accomplish this, many problems need to be addressed. The most relevant are detailed further.

Problems a Face Recognition System Needs to Address

In real life, faces appear under a variety of conditions. The most common ones include the following:

- *Pose*: Faces move in 3D space. When captured by a 2D camera, a large variety of 2D images corresponding to the same face can be obtained. Alternatively, one can use 3D scans, but these generally require the cooperation of the subject and are more expensive.
- *Illumination*: Different ambient lighting results in very distinct texture patterns of the face. One illumination will emphasize one type of face texture, while a different lighting will accentuate another. The shape of the face is also affected, because

different illumination angles will cause completely distinct cast shadows.

- *Expression*: Faces are a fundamental means by which we express emotions as well as other cues related to human communication. This is achieved by employing a large collection of the face muscles underlying the skin. With different movements come different expressions, each of which results in a distinct facial appearance.
- *Occlusions*: In most applications, partial occlusions are a common occurrence. These may be caused by clothing, glasses (including sunglasses), self-occlusions (such as a hand), clutter, etc. This means that when pictures of faces are taken, not all the information is always available. In fact, when 2D images are used, only a portion of the face is visible. The 3/4 view provides the most information, but even this orientation misses information because the face is not symmetric.
- *Imprecisely localized faces*: A less known problem of dealing with faces is that it cannot be precisely located or cannot be robustly delineated from an image or a video sequence. One reason is that it is generally impossible to determine where a face or facial feature starts and ends. To see this, an image of a face can be uploaded on favorite image software. Next, the inner corner of the left eye is zoomed in until the pixels become large squares. It can be seen at the pixel level that it is almost impossible to determine where the inner corner of left eye is. This problem makes the process of face detection difficult – even when we try to perform segmentation by hand.

Any successful algorithm for face recognition has to address some or all of these problems [3–6]. The great advances in recent years should be looked up with gratitude for there are algorithms now to partially solve each and every one of them. However, this may require tuning the approach to each application. There is a whole spectrum of techniques available to practitioners. The methodologies defined over the years vary considerably, from shape- to appearance-based recognition. The algorithms that are predominantly based on the shape of the face require extraction of the outline of the facial components, which has only been partially resolved recently [7, 8]. The appearance-based approach simplifies some of these requirements. In this alternative approach, one uses the brightness of the pixels defining the face as features for representing and recognizing images [9, 10]. Nonetheless, the correct definition of the approach still requires that the faces be warped to a "standard" shape, which involves the detection of some of the major facial components (e.g., eye centers) [3]. These approaches are summarized below.

Different Approaches to Face Recognition

The first step in understanding "Face Recognition" and designing systems that can do automatic recognition will undoubtedly be that of describing the face (see Anatomy of Face). The face is an articulate object capable of amazing transformations. While the underlying bone structure defines who we are – our identity and part of our heritage – the muscles overlaying it shape our personality. Muscles are also fundamental for the recognition of emotions and other communicative cues, although recent results [11, 12] demonstrate that other factors are also in the play and ought to be considered.

After the face anatomy has been studied, understanding how to acquire and design automatic systems for face recognition is needed, including a review of "Face Sample Quality". Sometimes it is taken for granted that the quality of the images we capture is the optimal for computational analysis, which is generally not the case.

After the basic components of the face and face recognition systems have been introduced, the processes of acquisition to recognition using algorithmical components are discussed. The first step is to determine the location of the face or faces in the image or video sequence (see Face Detection). Following this is the aspect of "Face Tracking," which is about how to track the motion of the face within a video sequence without the need to detect in each individual frame. Face detection and face tracking approaches are important because they either outperform the rest or because their mathematical formulation makes them appropriate in a large number of face recognition applications.

However, the processes of face detection and tracking do not generally suffice. If the problems of face recognition are to be appropriately addressed, faces are to be aligned before recognition – although this process can also be combined with that of identification

(▶ Face Alignment) and derive computer algorithms for aligning and warping faces ▶ Face Warping by means of previously devised models of faces (▶ Deformable Models).

The next aspect in face recognition is "Face Variations," including problems of illumination, pose, and expression. As illumination influences the acquisition of face images and their subsequent recognition, it is important to know how to do recognition under varying illumination. Pose variations include approaches on how to model faces seen from different points of view. Expression variations refer to how we can design algorithms for the recognition of emotions and other facial expressions. Of particular note are the applications in human–computer interaction.

After all the different types of features that can be used to represent, model, and recognize faces have been introduced, the actual problem of classification is discussed, which includes recognition from shape and appearance, local and global components, video, 3D range data, near and thermal infrared, and sketch. Each of these methods has advantages and disadvantages that make them appropriate in some scenarios but not in others and form an important part of how face recognition is performed. A combination of approaches is under study [13] and more of them might be seen as technology improves and costs decrease.

Face recognition algorithms cannot be tested in the absence of well-defined databases, which are very important while deriving face recognition systems and we need to understand how different algorithms and systems compare with each other.

Finally, there is a need to understand progress in skin color modeling and skin texture. These are related topics to those described above and each of them can benefit from one another.

Summary

The essay discusses what the face is, how it varies, and how it can be modeled and recognized using computer algorithms: the face and its variations; algorithms for detection, tracking, and modeling; major approaches for recognition; and databases, evaluation protocols and alternative mechanisms of modeling and recognition have been discussed elsewhere in the encyclopedia.

Related Entries

▶ Anatomy of Face
▶ And-Or Graph Model for Face Representation, Sketching and Aging
▶ Biometrics, Overview
▶ Deformable Models
▶ Face Alignment
▶ Face Databases and Evaluation
▶ Face Localization
▶ Face Pose Analysis
▶ Face Recognition, Component-Based
▶ Face Recognition, 3D-Based
▶ Face Recognition, Near Infrared Based
▶ Face Recognition, Overview
▶ Face Recognition, Shape vs. Appearance-Based
▶ Face Recognition, Thermal Infrared Based
▶ Face Recognition, Video-based
▶ Face Sample Quality
▶ Face Tracking
▶ Face Variation
▶ Facial Expression Recognition
▶ Mis-alignment Analysis
▶ Skin Color Modeling

References

1. Jain, A.K.: Technology: Biometric recognition. Nature **449**, 38–40 (2007)
2. Teoh, A.B.J., Ngo, D., Goh, A.: Personalised cryptographic key generation based on facehashing. Comput. Secur. **23**(7), 606–614 (2004)
3. Martinez, A.M.: Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class. IEEE Trans. Pattern Anal. Mach. Intell. **24**(6), 748–763 (2002)
4. M. Yang, D.J., Kriegman, N.A.: Detecting faces in images: A survey. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 34–58 (2002)
5. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Comput. Surv. **35**, 399–458 (2003)
6. Gross, R., Matthews, I., Baker, S.: Appearance-based face recognition and lightfields. IEEE Trans. Pattern Anal. Mach. Intell. **26**(4), 449–465 (2004)
7. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 681–685 (2001)
8. Ding, L., Martinez, A.: Precise detailed detection of faces and facial features. In: Proceedings of IEEE Computer Vision and Pattern Recognition, Anchorage, AK, 23 June 2008

9. Sirovich, L., Kirby, M.: A lowdimensional procedure for the characterization of human faces. J. Opt. Soc. Am. A **4**(3), 519–524 (1986)
10. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
11. Neth, D., Martinez, A.M.: Emotion perception in emotionless face images suggests a normbased representation. J. Vis. **9**(1), 1–11 (2009)
12. Zebrowitz, L.A.: Reading faces: window to the soul? Westview Press, Boulder, CO (1997)
13. Chang, K.I., Bowyer, K.W., Flynn, P.J.: An evaluation of multimodal 2d + 3d face biometrics. IEEE Trans. Pattern Anal. Mach. Intell. **27**(4), 619–624 (2005)

Face Recognition, Thermal

George Bebis
Department of Computer Science and Engineering, University of Nevada, Reno, NV, USA

Synonym

Face Recognition, Thermal Infrared

Definition

The human face emits thermal radiation which can be sensed by imaging sensors (i.e., thermal cameras) that are sensitive in the thermal infrared (IR) band of the ► electromagnetic (EM) spectrum. Temperature variations on the surface of the face produce a heat pattern, called a ► thermogram, which can be visualized as a 2D image (i.e., thermal image). Due to the presence of highly distinctive and permanent physiological characteristics under the facial skin, thermograms contain important information which can be exploited for face recognition.

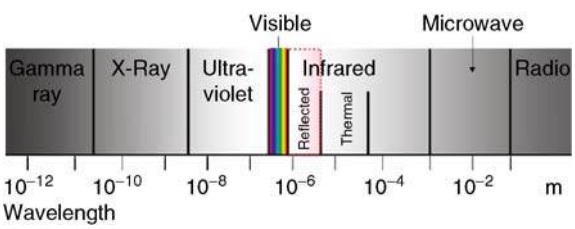
Introduction

Considerable progress has been made in face recognition over the last decade [1], however, face recognition technology is not accurate or robust enough to be deployed in uncontrolled environments, for example, protecting high value assets (e.g. perimeter of government buildings) from asymmetric (i.e., terrorist) threats. Human facial signatures vary significantly across races in the visible band. This variability, coupled with dynamic lighting conditions, presents a formidable problem. For instance, face recognition under very low lighting is almost impossible using visible imagery. Reducing light variability through the use of an artificial illuminator is rather awkward in the visible band because it may be distracting to the eyes of the people in the scene and reveals the existence of the surveillance system.

Thermal IR imagery offers a promising alternative to visible imagery for handling variations in face appearance due to illumination changes [2], facial expression [3, 4], and face pose [4] more successfully. In particular, thermal IR imagery is nearly invariant to changes in ambient illumination [2, 3], and provides a capability for the identification under all lighting conditions including total darkness [4]. Therefore, while face recognition systems in the visible spectrum opt for pure algorithmic solutions into inherent phenomenology problems, systems employing thermal IR imagery have the potential to offer simpler and more robust solutions, improving performance in uncontrolled environments and deliberate attempts to obscure identity [5].

Thermal IR Spectrum

Imaging sensors sensitive in the visible spectrum respond to ► electromagnetic radiation in the range (0.4–0.7 μ), while sensors sensitive in the IR spectrum respond to electromagnetic radiation in the range 0.7–14.0 μ (see Fig. 1). In general, the IR spectrum can be divided into two primary bands: the *reflected* IR and the



Face Recognition, Thermal. Figure 1 The electromagnetic (EM) spectrum.

thermal IR bands. The reflected IR band (0.7–2.4 μ) corresponds to reflected solar radiation and contains no information about the thermal properties of objects. It can be divided into two subbands: the near-ir (NIR) (0.7–0.9 μ) and the short-wave infrared (SWIR) (0.9–2.4 μ) bands.

The thermal IR band (2.4–14.0 μ) corresponds to thermal radiation emitted from objects. Temperature variations on the surface of an object produce a heat pattern, called *thermogram*, which can be visualized as a 2D image (i.e., thermal image). The amount of emitted radiation depends both on the temperature and the emissivity of the objects [6]. The thermal IR band can be divided into two subbands: the mid-wave infrared (MWIR) with range 3.0–5.0 μ and long-wave infrared (LWIR) with range 8.0–14.0 μ. It should be mentioned that there are strong atmospheric absorption bands at 2.4–3.0 μ and at 5.0–8.0 μ.

Due to the presence of highly distinctive and permanent physiological patterns under the facial skin (i.e., vein and tissue structure) [7], thermograms contain important information that can be exploited for face recognition. The human face emits thermal radiation both in the MWIR and LWIR bands of the thermal IR spectrum. However, thermal emissions of the skin are much higher in the LWIR band than in the MWIR band. As a result, face images have a much lower within-class variation in the LWIR spectrum. To analyze a thermal image, ► radiometric calibration is required. This is a process which achieves a direct relation between the value at a pixel of the thermal image and the absolute amount of thermal emission from the corresponding physical scene element. The goal is to standardize thermal IR images, independently of environmental conditions, cameras, and passage of time [3].

Recognition in the Thermal IR

Face recognition in the visible spectrum exploits the reflectance characteristics of the human face. As a result, changes in ambient illumination might degrade recognition performance. Face recognition in the thermal IR spectrum exploits physiological characteristics of the face by considering the thermal energy emitted from the face rather than the light reflected. Therefore, face recognition in the thermal infrared (IR) spectrum is nearly invariant to changes in ambient illumination. Moreover, it is less sensitive to scattering and absorption by smoke or dust while the tasks of face detection and localization can be simplified considerably due to the fact that background clutter is typically not visible. Early overviews of face identification in the thermal IR spectrum can be found in [8–10]. A recent review on face recognition methods, both in the visible and thermal IR bands, can be found in [11] while a general review on multispectral face recognition methods, with emphasis on thermal IR, can be found in [12].

The effectiveness of visible versus IR spectrum was compared in an early study using several recognition algorithms in [13]. Using a database of subjects without eyeglasses, varying facial expression, and allowing minor lighting changes, it was found that there are no significant performance differences between visible and IR recognition across all the algorithms tested. In later studies [3, 14, 15], several popular appearance-based face recognition methodologies were tested under various lighting conditions and facial expressions. Results from these studies indicate superior performance for thermal IR-based recognition comapred with that for visible-based recognition. These findings were confirmed in an operational scenario where images were captured both indoors and outdoors [16].

The effect of lighting, facial expression, and passage of time between the gallery and probe images were examined in [17]. Although IR-based recognition outperformed visible-based recognition assuming lighting and facial expression changes, it was found that IR-based recognition degrades when there is substantial passage of time between the gallery and probe images. In a related study [18], however, it was reported that both thermal IR and visible imagery degrade similarly with time passage. Improvements using fusion were reported in [16, 17]. Recognition using thermal IR was also shown to be less sensitive to changes in 3D head pose and facial expression in [4]. In [19], a statistical hypothesis pruning methodology was introduced for face recognition in thermal IR. First, each thermal IR face image was decomposed into spectral features using Gabor filters. Then, it was represented by a few parameters by modeling the marginal density of the Gabor filter coefficients using Bessel functions. Recognition was performed in the space of parameters of the Bessel functions.

Methodologically, the majority of the thermal IR face recognition methods reported in the earlier sections do not differ significantly from face recognition methods in

the visible band (i.e., appearance-based and feature-based). An exception is the method presented in [7] which explicitly exploits physiological information using the bioheat information present in thermal images. In particular, this method extracts the superficial blood vessel network which contains contour shapes quite characteristic of each individual. Matching is based on the branching points of the skeletonized vascular network. Using physiological features has the potential to improve thermal IR-based recognition by making recognition more robust to changes over time.

Limitations of Thermal IR

Despite its advantages, thermal IR has several drawbacks. First, it is sensitive to temperature changes in the surrounding environment. Currents of cold or warm air could influence the performance of systems using IR imagery. Second, it is sensitive to variations in the heat patterns of the face. Factors that could contribute to these variations include facial expressions (e.g. open mouth), physical conditions (e.g. lack of sleep, physical exercise), and psychological conditions (e.g. fear, stress, excitement). Third, thermal IR is opaque to glass. Glass blocks a large portion of thermal energy resulting in a loss of information near the eye region as shown on Fig. 2. Finally, radiometric calibration is required every time the environmental conditions change (e.g., moving the camera at a different location), a different camera is used (e.g., even if it is the same model), or data collections take place at different time intervals.

Fusion of Visible with Thermal IR Imagery

The benefits of fusing visible with thermal IR imagery have been documented in a number of studies including [3, 17, 20–23]. The idea is to combine the strengths of each spectral band to build more accurate and robust face recognition systems. For example, increased body temperature changes the thermal characteristics of the face, while there are not significant differences in the visible spectrum. Also, while eyeglasses completely occlude the eyes in the thermal IR spectrum, the problem is considerably less severe in the visible spectrum although visible imagery can suffer from highlights on the glasses under certain illumination conditions. A summary of the fusion strategy reported in [21–23] for improving face recognition performance in the presence of eyeglasses is as follows.

Objects made of glass act as a temperature screen, completely hiding the parts located behind them. In the case of subjects wearing eyeglasses, this poses some major difficulties since the eyes would be occluded completely due to the fact that eyeglasses block thermal energy (i.e., see Fig. 2). Experimental results, reported in [21–23], illustrate that face recognition performance in the thermal IR degrades seriously when eyeglasses are present in the probe image but not in the gallery image and vice versa. To address this limitation, fusion of thermal IR with visible imagery was employed in [21–23]. Two different fusion strategies were investigated: *pixel-based fusion* in the wavelet domain, and *feature-based fusion* in the eigenspace domain. In both cases, fusion was carried out using Genetic Algorithms (GAs) [24].

The Equinox database [25] was used for experimentation. The database contains frontal faces under the following scenarios: (1) three different light directions – frontal and lateral (right and left); (2) three facial expression – "frown," "surprise" and "smile"; (3) vocals pronunciation expressions – subjects were asked to pronounce several vocals from which three representative frames were chosen; and (4) presence of glasses – for subjects wearing glasses, all of these scenarios were repeated with and without glasses. For testing, the data were divided as follows: EG (expression frames with glasses, all illuminations), EnG (expression frames without glasses, all illuminations), EFG (expression frames with glasses, frontal illumination), ELG (expression frames with glasses, lateral illumination), EFnG (expression frames without glasses, frontal illumination), ELnG (expression frames without glasses, lateral illumination). The inclusion relations among these sets are as follows:

$$\begin{aligned} \mathrm{EG} &= \mathrm{ELG} \cup \mathrm{EFG}, \\ \mathrm{EnG} &= \mathrm{ELnG} \cup \mathrm{EFnG} \quad \text{and} \quad \mathrm{EG} \cap \mathrm{EnG} = \emptyset \end{aligned} \tag{1}$$

Recognition performance was measured by finding the percentage of the images in the test set, for which the top match is an image of the same person from the gallery. Figures 3 and 4 show the results obtained. Among the two fusion strategies tested, fusion in the wavelet domain yielded the best results. Nevertheless, fusion outperformed each modality alone in both cases.

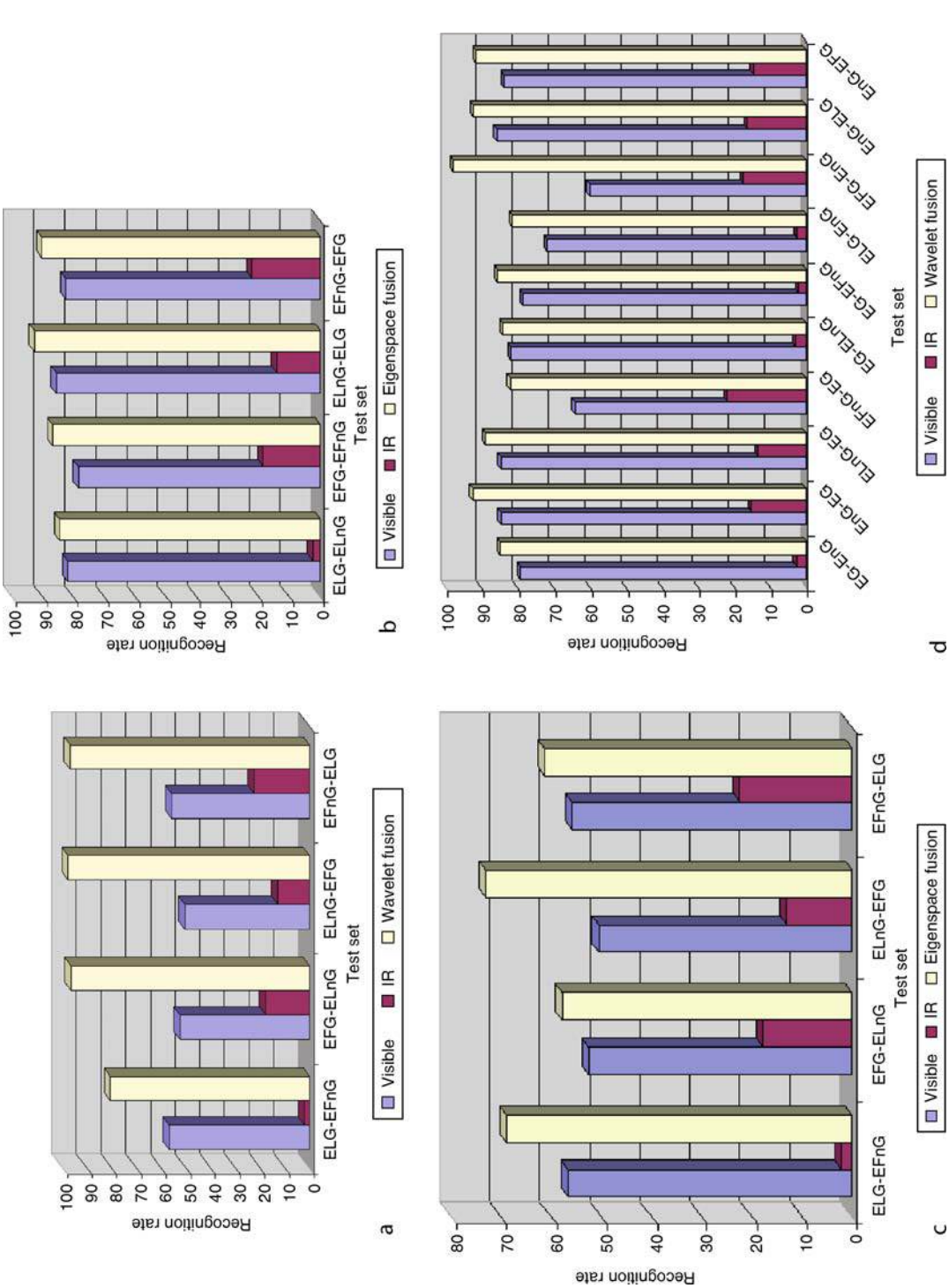
Face Recognition, Thermal. Figure 2 (**a, b**) Visible images; (**c, d**) thermal IR images. It should be observed that since thermal IR is opaque to glass, the presence of eyeglasses blocks the eyes completely.

Milestones

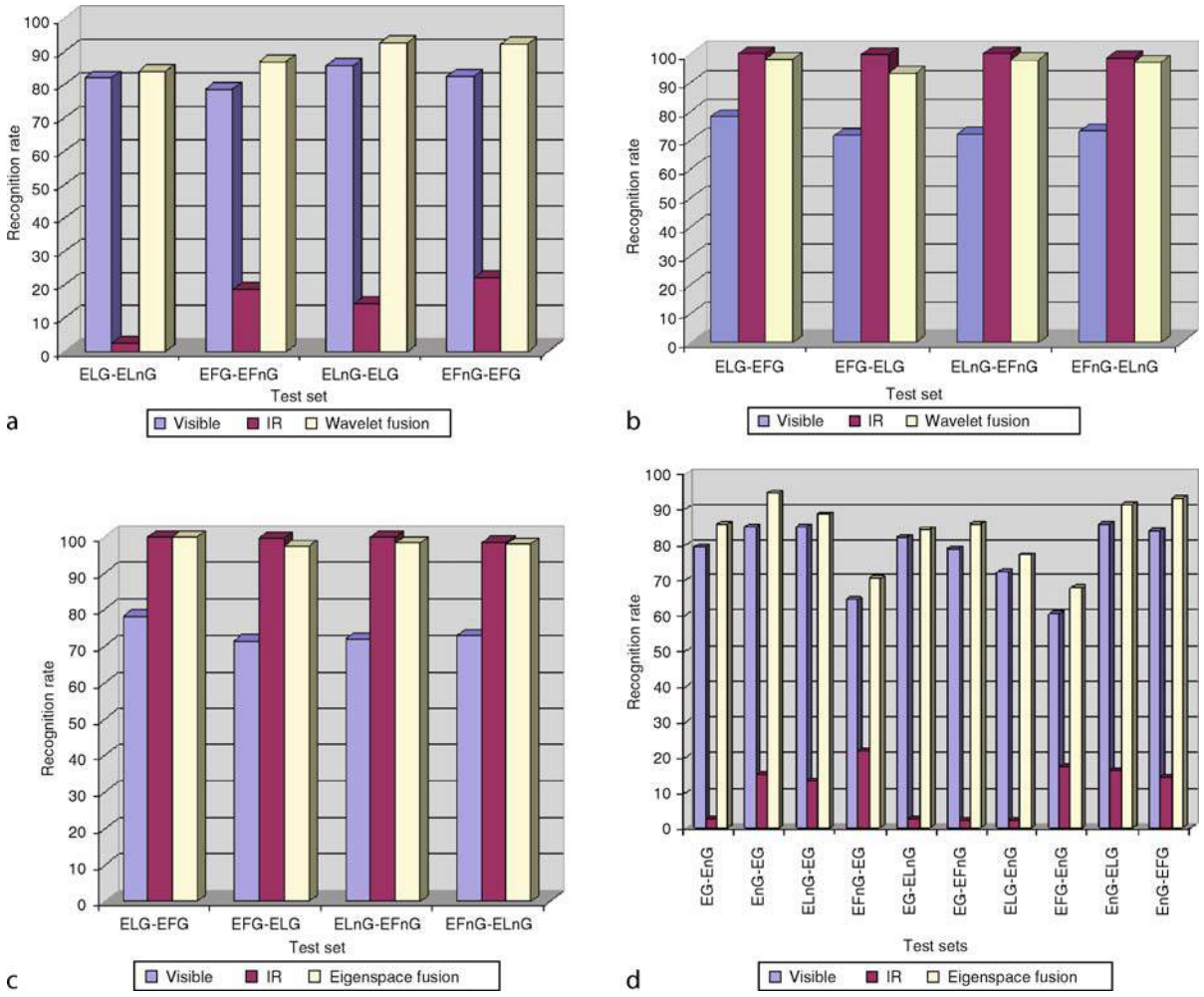
Despite its advantages, face recognition in the thermal IR spectrum has received relatively little attention compared to visible spectrum, mostly because of the following reasons:

- Higher cost of thermal sensors
- Lower image resolution
- Higher image noise
- Lack of widely available data sets

Advances in IR imaging technology and the availability of publicly available datasets, however, have facilitated experimentation with thermal imagery in the context of face recognition. While the difference in cost between visible and thermal imaging equipment is still large, the gap is closing rapidly as new uncooled microbolometer technologies enter the market. Companies, such as FLIR Systems (http://www.flir.com), offer a large variety of thermal cameras for different budgets. The issue of image noise can be addressed



Face Recognition, Thermal. Figure 3 Eyeglasses results in the wavelet domain: (**a**) same illumination conditions – eyeglasses are not present both in the gallery and probe sets; (**b**) eyeglasses are present both in the gallery and probe sets – illumination conditions are different; (**c**) eyeglasses are not present both in the gallery and probe sets – illumination conditions are different; (**d**) similar to (c) except that the gallery and probe sets contain multiple illuminations.



Face Recognition, Thermal. Figure 4 Eyeglasses results in the eigenspace domain: (**a**) same illumination conditions – eyeglasses are not present both in the gallery and probe sets; (**b**) eyeglasses are present both in the gallery and probe sets – illumination conditions are different; (**c**) eyeglasses are not present both in the gallery and probe sets – illumination conditions are different; (**d**) similar to (c) except that the gallery and probe sets contain multiple illuminations.

using powerful radiometric calibration procedures while the issue of image resolution can be addressed using super-resolution techniques or fusing infrared with visible imagery. For example, Equinox Corporation (http://www.equinoxsensors.com) has made available a system for real-time fusion of thermal IR with visible imagery with image co-registration correction.

In terms of data, things are rather limited compared to the plethora of face databases available in the visible spectrum [26, 27]. The most extensive IR facial database, that is publicly available, is the Equinox database [25]. This database was created by Equinox Corporation under DARPA's HumanID program. It includes coregistered visible/LWIR/MWIR/SWIR images and it is representative of unconstrained frontal imagery of people's faces in an indoor environment. Another, publicly available database, is the Notre Dame IR face database [28], which includes data captured at different sessions over time. Obviously, additional datasets are required to spark more research in this area, in particular, data depicting scenarios widely different from the imaging conditions during data acquisition (e.g., outdoor imagery). Introducing appearance variability due to various factors (e.g., metabolic activity) would be extremely useful in testing the robustness of thermal IR for face recognition.

Summary

While face recognition in the visible band performs satisfactorily under controlled conditions, thermal IR face recognition offers more advantages when there is no control over illumination or for detecting disguised faces. The passive nature of thermal IR systems lowers their complexity and improves their reliability. With dropping prices and technological advances, thermal IR is becoming more affordable and practical than before. Although thermal IR has many advantages, it suffers from several drawbacks including that it is sensitive to temperature changes and opaque to glass. A promising approach to deal with these issues is fusing visible with thermal IR imagery.

Related Entries

► Biometrics, Overview
► Face Recognition

References

1. Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A.: Face recognition: A literature survey. ACM Comput. Surveys **35**(4), 399–458, (2003)
2. Wolff, L., Socolinsky, D., Eveland, C.: Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery. In: IEEE Workshop on Computer Vision Beyond the Visible Spectrum. Hawaii (2001)
3. Socolinsky, D., Selinger, A., Neuheisel, J.: Face recognition with visible and thermal infrared. Comput. Vision Image Understand **91**, 72–114 (2003)
4. Friedrich, G., Yeshurun, Y.: Seeing people in the dark: face recognition in infrared images. In: International Workshop on Biologically Motivated Computer Vision, pp. 348–359 (2002)
5. Pavlidis, I., Symosek, P.: The imaging issue in an automatic face/disguise detection system. In: IEEE Workshop on Computer Vision Beyond the Visible Spectrum, pp. 15–24 (2000)
6. Siegel, R., Howell, J.: Thermal Radiation Heat Transfer, 3rd edn. Taylor & Francis, London (1992)
7. Buddharaju, P., Pavlidis, I., Tsiamyrtzis, P., Bazakos, M.: Physiology-based face recognition in the thermal infrared spectrum. IEEE Trans. Pattern Analy. Mach. Intell., **29**(4), 613–626 (2007)
8. Evans, D.: Infrared facial recognition technology being pushed toward emerging applications. In: Proceedings of SPIE, vol. 2962, pp. 276–286 (1997)
9. Prokoski, F., Riedel, R.: Infrared identifications of faces and body parts. In: Jain, A., Bolle, R., Pankanti, S. (eds.) BIOMETRICS: Personal Identification in Networked Society, pp. 191–212 (1999)
10. Prokoski, F.: History, current status, and future of infrared identification. In: IEEE Workshop on Computer Vision Beyond the Visible Spectrum. Hilton Head (2000)
11. Kong, S., Heo, J., Abidi, B., Paik, J., Abidi, M.: Recent advances in visual and infrared face recognition – a review. Comput. Vision Image Understand. **97**, 103–135 (2005)
12. Socolinsky, D.: Multispectral Face Recognition. In: HandBook of Biometrics, Jain, A., Flynn, P., and Ross, A. (eds.) pp. 293–313 (2008)
13. Wilder, J., Phillips, J., Jiang, C., Wiener, S.: Comparison of visible and infra-red imagery for face recognition. In: IEEE International Conference on Automatic Face and Gesture Recognition, pp. 182–187. Killington (1996)
14. Socolinsky, D., Wolff, L., Neuheisel, J., Eveland, C.: Illumination invariant face recognition using thermal infrared imagery. In: Computer Vision and Pattern Recognition Conference. Hawaii (2001)
15. Socolinsky, D., Selinger, A.: Comparative study of face recognition performance with visible and thermal infrared imagery. In: International Conference on Pattern Recognition, pp. 217–222 (2002)
16. Socolinsky, D., Selinger, A.: Thermal face recognition in an operational scenario. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2004)
17. Chen, X., Flynn, P., Bowyer, K.: Ir and visible light face recognition. Comput. Vision Image Understand. **99**, 332–358 (2005)
18. Socolinsky, D., Selinger, A.: Thermal face recognition over time. In: International Conference on Pattern Recognition (ICPR) (2004)
19. Srivastana, A., Liu, X.: Statistical hypothesis pruning for recognizing faces from infrared images. Image Vision Comput. **21**, 651–661 (2003)
20. Heo, J., Kong, S., Abidi, B., Abidi, M.: Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In: Workshop on Object Tracking and Classification Beyond the Visible Spectrum (2004)
21. Gyaourova, A., Bebis, G., Pavlidis, I.: Fusion of infrared and visible images for face recognition. In: European Computer Vision Conference (ECCV) (2004)
22. Singh, S., Gyaourova, A., Bebis, G., Pavlidis, I.: Infrared and visible image fusion for face recognition. In: SPIE Defense and Security Symposium (Biometric Technology for Human Identification) (2004)
23. Bebis, G., Gyaourova, A., Singh, S., Pavlidis, I.: Face recognition by fusing thermal infrared and visible imagery. Image Vision Comput. **24**(7), 727–742 (2006)
24. Goldberg, D.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison Wesley, Reading, MA, USA (1989)
25. Equinox corporation, ir face database. http://www.equinoxsensors.com/products/HID.html (last visited in June 2008)
26. Gross, R.: Face Databases. In: Li, S.Z., Jain, A.K. (eds) Handbook of Face Recognition, pp. 301–327 (2005)
27. Face recognition homepage. http://www.face-rec.org/ (last visited in June, 2008)
28. Computer vision laboratory, university of notre dame, biometrics database distribution. http://www.nd.edu/cvrl/ (last visited in June, 2008)

Face Recognition, Video-Based

Rama Chellappa[1], Gaurav Aggarwal[1],
S. Kevin Zhou[2]
[1]University of Maryland, College Park, USA
[2]Siemens Corporate Research, Princeton, NJ, USA

Synonyms

Face recognition from image sequences; Video-based face recognition

Definition

Video-based face recognition is the technique of establishing the identity of one or multiple persons present in a video, based on their facial characteristics. Given the input face video, a typical video-based face recognition approach combines the temporal characteristics of facial motion with appearance changes for recognition. This often involves ▶ temporal characterization of faces for recognition, building 3D model or a super-resolution image of the face, or simply learning the appearance variations from the multiple video frames. The ability to generalize across pose, illumination, expression, etc. depends on the choice of combination. Video-based face recognition is particularly useful in surveillance scenarios in which it may not be possible to capture a single good frame as required by most still image based methods.

Introduction

Face recognition is one of the most successful applications in the vast amount of research on image analysis and understanding [1]. The fact that face recognition can be performed at a distance without subject's cooperation or knowledge makes it particularly attractive as compared to more reliable biometrics like fingerprints and iris or retinal scans. Traditionally face recognition has been limited to still images. Though great leaps have been made in recognizing faces from still images, more needs to be done to achieve the goal of recognizing faces in uncontrolled scenarios. Still image based approaches often struggle to truly generalize across variations in pose, expression, illumination, etc., leading to a not so satisfactory performance on real images. The advent of inexpensive cameras and increased processing power has made it possible to capture and store videos in real time. Videos have the advantage of providing more information in the form of multiple frames making it relatively easier to generalize across variations that have been difficult with still images. Moreover, video makes it easier to track (or segment) faces which can then be fed into a recognition system. Importantly, psychological evidence indicates that dynamic information contributes to face recognition especially under nonoptimal viewing conditions [2]. These reasons form the basis of the recent interest in using videos for recognizing faces [3–5]. Though video provides extra information, the video feeds are almost always uncontrolled making it challenging to track and hence recognize faces.

Operation of a Video-based Face Recognition System

A typical Video-based Face Recognition (VFR) system operates by acquiring video feeds from one or multiple cameras, tracking and segmenting faces from the input feed(s), extracting representations to characterize the identity of the face(s) in the video, and then comparing them with the enrolled representations of subjects in the database. This constitutes the test phase of the system. During the enrollment (or training) phase, a similar sequence of steps is followed using one or multiple video feeds per identity and the corresponding composite representations are stored in the database. VFR approaches differ in the representation that is used to characterize the moving faces. An ideal VFR system performs these operations automatically without any human intervention. Though potentially a VFR system can operate in either verification mode (one-to-one matching) or identification mode (one-to-many), the real application of such a system lies in identifying subjects using surveillance cameras (say on an airport) without their knowledge. Therefore, a typical VFR system will often operate in what is known as watch list mode [6]. The watch list problem is a generalization of both identification and verification problems in which the system only attempts the identification of individuals on the watch list. The performance in this mode is measured using both identification rate and false alarm rate.

Challenges for Video-based Face Recognition Systems

Effective utilization/fusion of the information (both spatial and temporal) present in a video to achieve better generalization (for each subject) and discriminability (across different subjects) for improved identification is one of the biggest challenges faced by a VFR system. The fusion schemes can range from simple selection of good frames (which are then used for recognition in a still-image based recognition framework) to estimation of the full 3D structure of a face which can then be used to generalize across pose, illumination, etc. The choice may depend primarily on the operational requirements of the system. For example, in a surveillance setting, the resolution of the faces may be too small for reliable shape estimation. The choice also limits the recognition capability of the system. A simple good frame selection scheme will not have the capability to generalize appearance across pose variations and thus requires the test video to have some pose overlap with the gallery videos. Effective modeling of subject-specific facial characteristics from video data can only be achieved if the changes in facial appearance during the course of the video are appropriately attributed to different factors like pose changes, lighting, expression variations, etc. Unlike still image based scenarios, these variations are inherent in a VFR setting and must be accounted for to reap the benefits of extra information provided by the video data. In addition, due to the nature of the input data, VFR is often addressed in conjunction with tracking problem which is a challenging problem by itself. In fact, more often than not, tracking accuracy depends on the knowledge of reliable appearance model (depends on the identity provided by the recognition module) while recognition result is dependent on the localization accuracy of the face region in input video.

Examples of Video-based Face Recognition Algorithms

Given the potential advantages video provides for the task of face recognition, relatively little work has been done to recognize faces in videos. The challenges in modeling moving faces along with the unavailability of large standard datasets have hindered the progress of research on VFR algorithms. Table 1 gives a snapshot of a few existing VFR algorithms. As clear from the table, all the approaches have been tested on a very small sized (often private) datasets. The following discussion describes them in detail.

1. *Simultaneous Tracking and Recognition of Faces:* Traditional tracking-then-recognition approaches resolve uncertainties in tracking and recognition sequentially and separately, which often involves difficult choices (like criteria to select good frames and

Face Recognition, Video-Based. Table 1 A snapshot of a few existing video-based face recognition algorithms

Algorithm	Short description	Experimental evaluation
Probabilistic recognition of human faces from video [7]	Simultaneous tracking-and-recognition using a time series state space model and sequential importance sampling	Private: 12 subjects, NIST: 30 subjects, MoBo [8]: 25 subjects
Video-based face recognition using probabilistic appearance manifolds [9]	Face modeled using a low-dimensional appearance manifold, approximated by piecewise linear subspaces	*Honda-UCSD* dataset: 20 subjects (52 videos)
Face verification through tracking facial features [10]	Tracks facial features defined on a grid with Gabor attributes using SIS algorithm	*Li* dataset: 19 subjects (2 sequences each)
Video-based face recognition using adaptive hidden markov models [11]	Statistics of training videos, and their temporal dynamics learnt by an HMM	Private: 12 subjects, Mobo [8]: 25 subjects
A system identification approach for video-based face recognition [12]	Face modeled as a linear dynamical system using ARMA model	*Honda-UCSD* dataset [9]: 30 subjects, *Li dataset* [10]: 19 subjects

estimation of registration parameters). Zhou et al. [7] avoid these issues while resolving uncertainties in tracking and recognition simultaneously in a unified probabilistic framework. The temporal information present in the video is fused using a time-series state-space model to characterize the evolving kinematics and identity. The three basic components of the model are as follows.

- A motion equation governing the kinematic behavior of the tracking motion vector. In its most general form, the motion equation can be written as

$$\theta_t = g(\theta_{t-1}, u_t); \quad t \geq 1, \tag{1}$$

  where u_t is the noise that determines the transition probability $p(\theta_t|\theta_{t-1})$. The function $g(.,.)$ characterizes the evolving motion. It can either be a function learned offline or given a priori. Choice of θ_t is application dependent.
- An identity equation governing the temporal evolution of the identity variable.

$$n_t = n_{t-1}; \quad t \geq 1, \tag{2}$$

- An observation equation establishing the link between the motion vector and the identity variable.

$$\tau_{\theta_t}\{z_t\} = I_{n_t} + v_t; \quad t \geq 1, \tag{3}$$

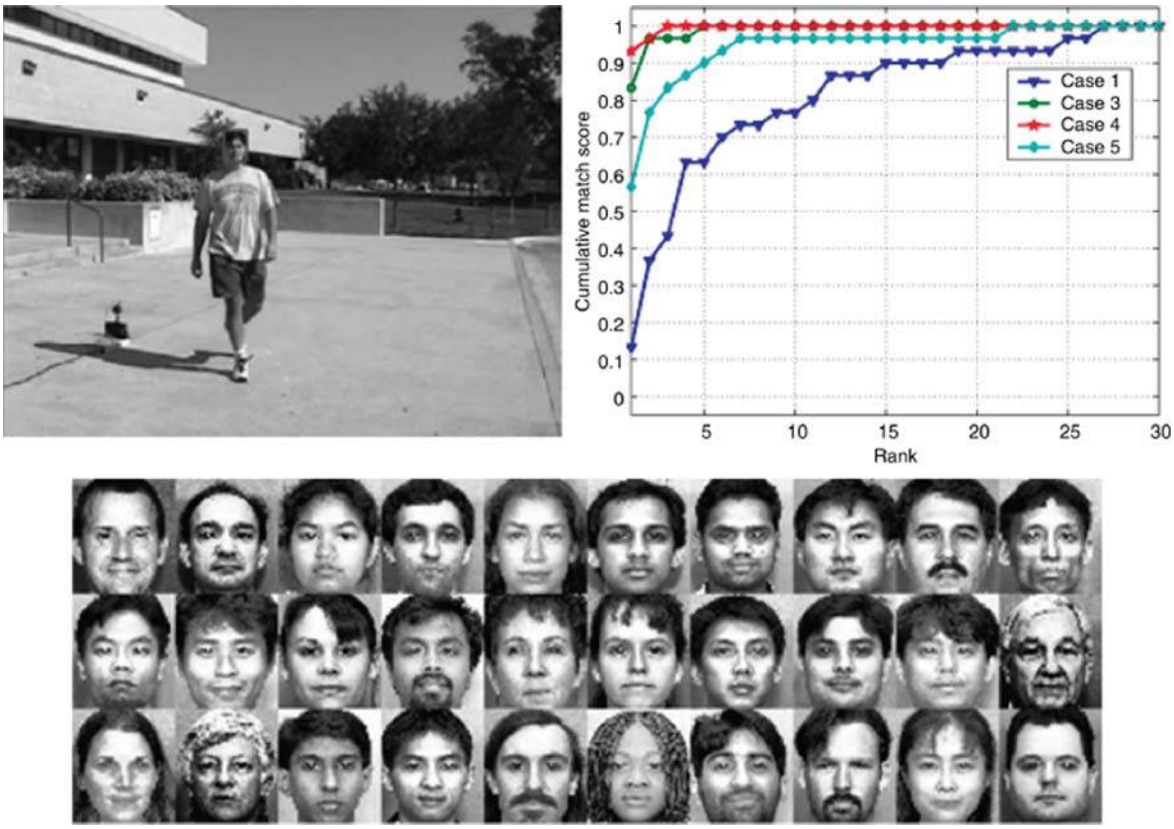
  where v_t is the observation noise that determines the observation likelihood $p(z_t|n_t,\theta_t)$ and $\tau_{\theta_t}\{z_t\}$ transforms the observation z_t to the chosen feature space.

Under the assumption of statistical independence between all noise variables and prior knowledge of the distributions $p(\theta_0|z_0)$ and $p(n_0|z_0)$, (1) and (2) can be combined as follows.

$$p(x_t|x_{t-1}) = p(n_t|n_{t-1})p(\theta_t|\theta_{t-1}), \tag{4}$$

where $x_t = (\theta_t, n_t)$. Given a video sequence, the goal is to estimate the posterior probability $p(n_t|z_{0:t})$. The posterior probability is calculated using Sequential Importance Sampling (SIS) [13]. Using the SIS technique, the joint probability distribution of the motion vector and the identity variable is estimated at each time instant and then propagated to the next time instant as governed by the motion and identity equations. The marginal distribution of the identity variable is estimated to provide the desired identity result. Fig. 1 shows the performance of the approach on a NIST dataset consisting of 30 persons gallery.

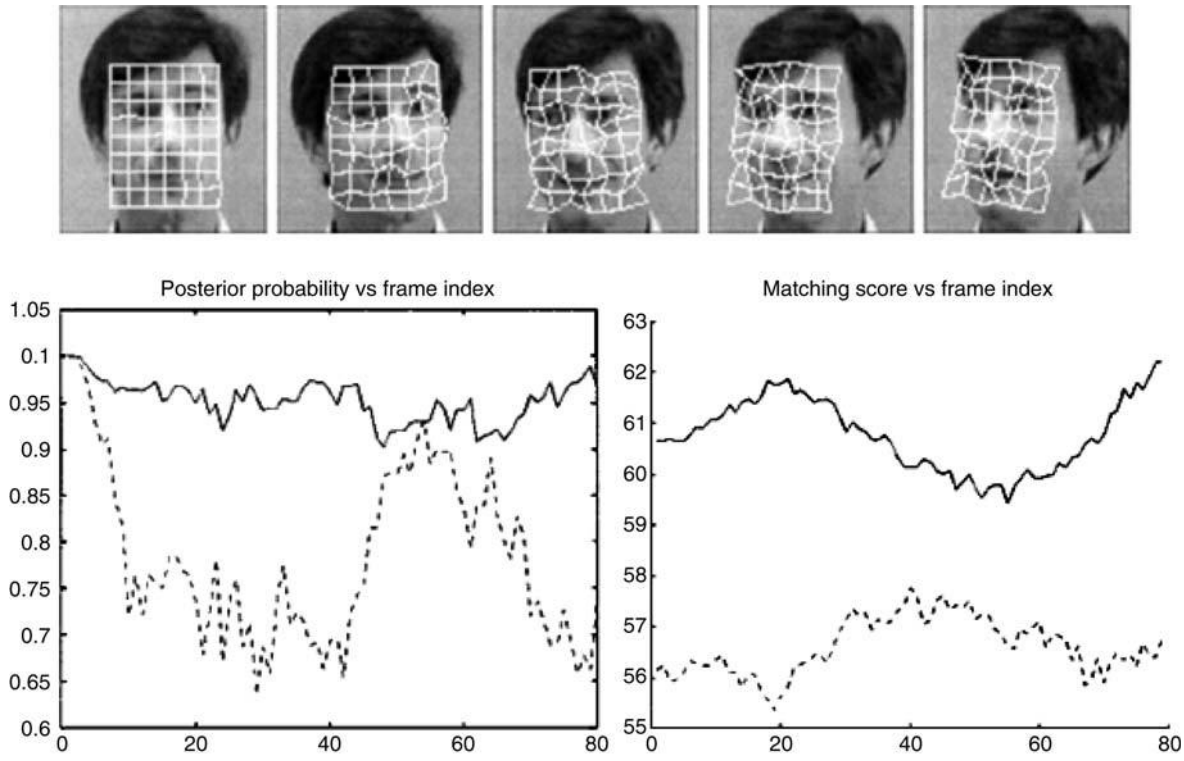
2. *Probabilistic Appearance Manifolds for VFR:* Similar to [7], Lee et al. [9] propose a VFR algorithm that performs modeling, tracking and recognition in one integrated framework. This is accomplished using a probabilistic appearance manifold based representation that is utilized simultaneously by both tracking and recognition modules. The recognition module uses tracker's output (the location of the face in the current frame) to update the current internal appearance model that is in turn used by the tracker.
   Each face is characterized using a collection of linear subspaces in the image space which is constructed by clustering the exemplars from the input face videos. Each cluster often contains face images with similar poses and is represented using a PCA subspace. The collection of linear subspaces is further characterized using a transition matrix that captures the probabilities of moving from one pose subspace to another between two consecutive frames. The transition matrix is used to combine facial appearance with temporal coherency of pose variations to perform recognition. The approach has been tested on 52 video sequences of 20 different subjects.
3. *2D Feature Graph based Approach:* Li and Chellappa [10] propose a 2D feature-graph based approach for VFR in which the intensity model is replaced by a feature-graph using Gabor transform. The feature-graph approach is more robust to the variations in illumination and pose but possibly requires slightly higher-resolution videos. The tracking problem is formulated as a Bayesian inference problem for which Markov Chain Monte Carlo (MCMC) techniques are employed to obtain an empirical solution. A reparameterization is used to facilitate empirical estimation and to allow verification to be addressed simultaneously along with tracking. The facial features to be tracked are defined on a grid with Gabor attributes (Fig. 2). The motion of facial feature points is modeled as a global two-dimensional affine transformation (to account for head motion) plus a local deformation to account for the residual due to expression changes and modeling errors. The global motion is estimated by importance sampling while the residual motion is handled by incorporating local deformation into the likelihood measurement.



Face Recognition, Video-Based. **Figure 1** Sequential tracking and recognition [7]. Top left: A probe video from the NIST dataset; Top right: Recognition performance under different models; Bottom: Gallery set.

The temporal evolution of the jet positions is modeled as a dynamic system, where tracking is solved by analyzing this system, which in general, is non-Gaussian and nonlinear. Note that tracking is solved by analyzing the dynamic system governing the evolution of the changes in affine parameters. This reparameterization originates from a simple Taylor expansion. However, its novelty comes from the fact that one can choose different initial states for different purposes. If the initial state corresponds to a feature set from the first frame of a sequence, then the reparameterization is suitable for pure tracking. However, if the initial state represents some template from a candidate list, then the reparameterization is naturally good for tracking-for-verification. When a template and the sequence belong to the same person, tracking results should reflect a coherent motion induced by the same underlying shape. On the other hand, a more random motion pattern will often be observed when the template and the sequence belong to different persons. Thus, with different templates, such a tracker allows verification to be addressed simultaneously with tracking. The motion coherence in a shape is evaluated by calculating the posterior probabilities from the estimated densities on a region centered on the mean shape. Fig. 2 shows the tracking and verification results using this approach.

4. *Hidden Markov Models for VFR:* Li and Chen [11] propose adaptive Hidden Markov Models (HMM) to recognize faces in videos. During training, a separate HMM is learnt for each subject in the gallery to characterize appearance statistics and temporal dynamics of the facial motion. Recognition is performed by analyzing the test video by HMMs corresponding to subjects in the gallery. During the recognition process, test video sequences are used to update the gallery models in an unsupervised fashion based on the recognition result. The approach has been tested on two datasets with 21 and 24 subjects respectively.
5. *3D Model based Approach:* As opposed to most VFR approaches which model face as a 2D object, the

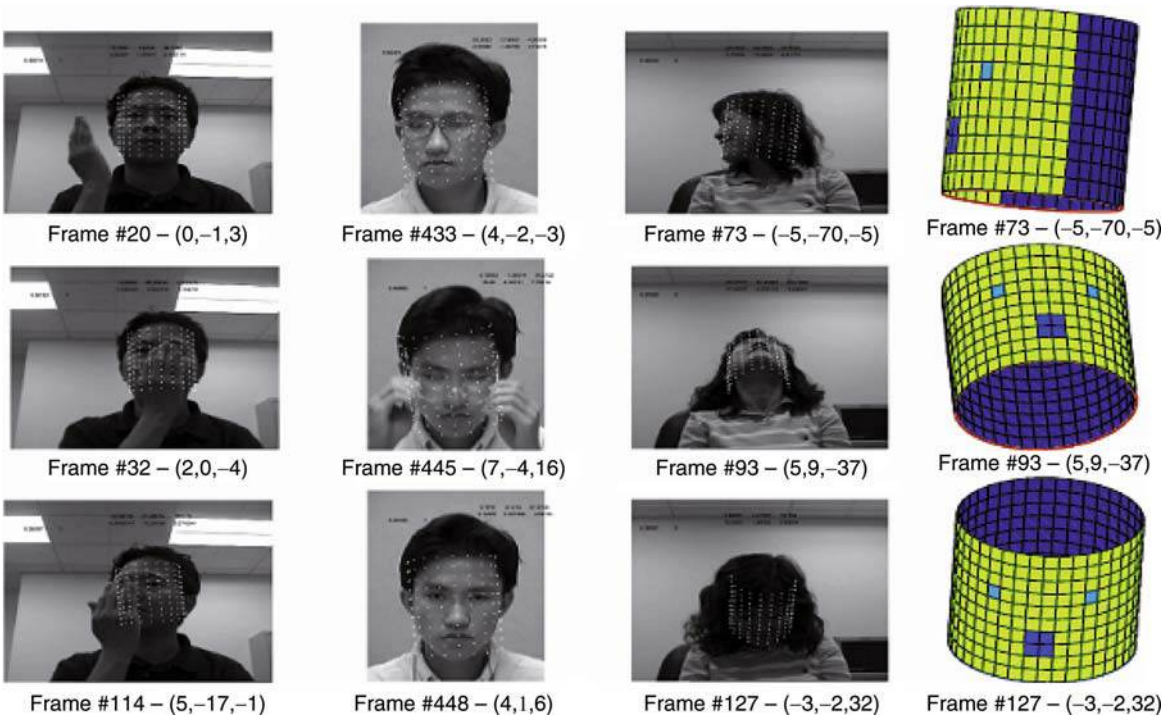


Face Recognition, Video-Based. **Figure 2** 2D feature-based approach [10]. Top: Tracking result; Bottom left: Posterior probabilities for the true (solid) and an impostor (dashed) hypothesis; Bottom right: Matching scores for the true (solid) and impostor hypothesis.

algorithm proposed in [14] estimates the 3D configuration of the head in each frame of the video. The 3D configuration consists of three translation parameters and three orientation parameters which correspond to the yaw, roll and pitch of the face. The approach combines the structural advantages of geometric modeling with the statistical benefits of a particle-filter based inference. The face is modeled as the curved surface of a cylinder which is free to translate and rotate in an unprescribed manner. The geometric modeling takes care of pose changes and self-occlusions while the statistical modeling handles unexpected occlusions and illumination variations during the course of the video. The recovered 3D facial pose information can be used to perform pose normalization which makes it very useful for the tasks of face modeling, face recognition, expression analysis, etc.

The estimation of 3D pose of a face in each frame of a video is posed as a dynamic state estimation problem. Particle filtering is used for estimating the unknown dynamic state of a system from a collection of noisy observations. Such an approach involves two components: 1) a state transition model to govern the motion of the face, and 2) an observation model to map the input video frames to the state (3D configuration). Figure 3 shows the tracking results for a few video frame. The estimated pose is shown in the form of a overlaid cylindrical grid. The accuracy in recovering 3D facial pose information makes it viable to perform VFR without any need for pose overlap between the gallery and test video. Recognition experiments are performed on videos with nonoverlapping poses. For each face, a texture mapped cylindrical representation is built using the recovered facial pose information, which is used for matching. The approach has been tested on a small dataset consisting of 10 subjects.

6. *Shape-Illumination Manifold for VFR:* In [15], Arandjelovic and Cipolla propose a generic shape-illumination manifold based approach to recognize faces in videos. Assuming the intensity of each pixel in an image to be a linear function of the corresponding albedo, the difference in two



Face Recognition, Video-Based. **Figure 3** 3D model-based approach [14]. The last column shows the pose for the frames in the third column.

logarithm-transformed images of the same subject in the same pose, depends only on 3D shape of the face and the illumination conditions in the input images. As the pose of the subject varies, the difference-of-log vectors describe manifold called as shape-illumination manifold in the corresponding vector space. Assuming shape variations across faces of different subjects to be small, a generic shape-illumination manifold (gSIM) can be learnt from a training corpus.

Given a test video for recognition, it is first re-illuminated in the illumination condition of each gallery video. Re-illumination involves a genetic algorithm (GA) based pose matching across the two face videos. For re-illumination, each frame of the test video is recreated using a weighted linear combination of K nearest neighbor frames of the gallery video as discovered by the pose matching module. This is followed by generation of difference-of-log vectors between each corresponding frame of the original and re-illuminated test videos. If the gallery and test video belong to the same subject, the difference-of-log vectors depend only on shape and illumination conditions. On the other hand, if the two videos come from different subjects, the vectors also depend on the differences in albedo maps of the two subjects. Finally, the similarity score is obtained by computing the likelihood of these postulated shape-illumination manifold samples under the learnt gSIM. The approach provides near perfect recognition rates on three different datasets consisting of 100, 60 and 11 subjects respectively.

7. *System Identification Approach:* Aggarwal et al. [12] pose VFR as a dynamical system identification problem. A moving face is modeled as a linear dynamical system whose appearance changes with pose. Each frame of the video is assumed to be the output of the dynamical system particular to the subject. Autoregressive and Moving Average (ARMA) model is used to represent such a system as follows

$$\begin{aligned} x(t+1) &= Ax(t) + v(t) \\ y(t) &= Cx(t) + \omega(t) \end{aligned} \tag{5}$$

Here $y(t)$ is the noisy observation of input $I(t)$ at time t, such that $y(t) = I(t) + \omega(t)$. I(t) is the appearance of face at time t and $x(t)$ is the hidden state that characterizes the pose, expression, etc. of

the face at time t. A and C are the system matrices characterizing the system, and $v(t)$ is an IID realization from some unknown density $q(.)$. Given a sequence of video frames, Aggarwal et al. [12] use a closed-form solution to estimate A and C. The similarity between a gallery and a probe video is measured using metrics based on subspace angles obtained from the estimated system matrices. The metrics used include Martin, gap, and Frobenius distance, all of which give similar recognition performance. The approach does well on the two datasets tested in [12]. Over 90% recognition rate is achieved (15/16 for the *Li* dataset [10] and 27/30 for the *UCSD/Honda* dataset [9]). The performance is quite promising given the extent of the pose and expression variations in the video sequences.

Summary and Discussion

There is little doubt that presence of multiple video frames allows for better generalization of person-specific facial characteristics over what can be achieved from a single image. In addition, VFR provides operational advantages over traditional still image based face recognition systems. Most existing VFR approaches have only been tested on independently captured very small datasets. Large standard datasets are required for better evaluation and comparison of various approaches.

Related Entries

- ▶ Face Recognition Overview
- ▶ Face Recognition Systems
- ▶ Face Tracking

References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Comput. Surv. **35**(4), 399–458 (2003)
2. O'Toole, A.J., Roark, A., Abdi, H.: Recognizing moving faces: A psychological and neural synthesis. Trends Cogn. Sci. **6**, 261–266 (2002)
3. Hadid, A., Pietikainen, M.: An experimental investigation about the integration of facial dynamics in video-based face recognition. Electronic Lett. Comput. Vis. Image Anal. **5**(1), 1–13 (2005)
4. Ekenel, H., Pnevmatikakis, A.: Video-based face recognition evaluation in the chil project - run 1. In: Proceedings of the seventh International Conference on Automatic Face and Gesture Recognition, pp. 85–90 (2006)
5. Gorodnichy , D. O. (Editor): Face processing in video sequences. Image and Vis. Comput. **24**(6), 551–648 (2006)
6. Grother, P., Micheals, R., Phillips, P.: Face recognition vendor test 2002 performance metrics. In: Proceedings of fourth International Conference on Audio and Video-Based Biometric Person Authentication, pp. 937–945 (2003)
7. Zhou, S., Kruger, V., Chellappa, R.: Probabilistic recognition of human faces from video. Comput. Vis. Image Underst. **91**(1–2), 214–245 (2003)
8. Gross, R., Shi, J.: The cmu motion of body (mobo) database. Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (2001)
9. Lee, K.C., Ho, J., Yang, M.H., Kriegman, D.: Visual tracking and recognition using probabilistic appearance manifolds. Comput. Vis. Image Underst. **99**(3), 303–331 (2005)
10. Li, B., Chellappa, R.: Face verification through tracking facial features. J. Opt. Soc. Am. A **18**(12), 2969–2981 (2001)
11. Liu, X., Chen, T.: Video-based face recognition using adaptive hidden markov models. In: Proceedings of International Conference on Computer Vision and Pattern Recognition, pp. 340–345 (2003)
12. Aggarwal, G., Roy-Chowdhury, A.K., Chellappa, R.: A system identification approach for video-based face recognition. In: Proceedings of International Conference on Pattern Recognition, pp. 175–178 (2004)
13. Liu, J.S.: Monte carlo strategies in scientific computing. Springer (2002)
14. Aggarwal, G., Veeraraghavan, A., Chellappa, R.: 3d facial pose tracking in uncalibrated videos. In: Proceedings of International Conference on Pattern Recognition and Machine Intelligence, pp. 515–520 (2005)
15. Arandjelovic, O., Cipolla, R.: Face recognition for video using the general shape-illumination manifold. In: Proceedings of European Conference on Computer Vision, pp. 27–40 (2006)

Face Reconstruction

- ▶ Forensic Evidence of Face

Face Registration

- ▶ Face Alignment

Face Sample Quality

Kui Jia[1], Shaogang Gong[2]
[1]Shenzhen Institute of Advanced Integration Technology, CAS/CUHK, Shenzhen, People's Republic of China
[2]Queen Mary, University of London, London, UK

Synonyms

Face sample standardization; Face sample utility

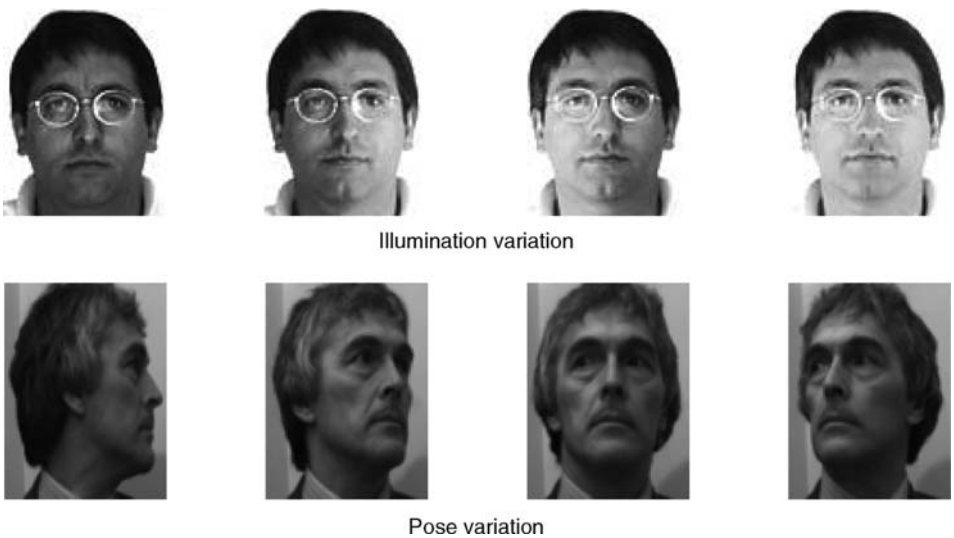
Definition

Face is a human biometric attribute that can be used to establish the identity of a person. A face-based biometric system operates by capturing probe face samples and comparing them against gallery face templates. The intrinsic characteristic of captured face samples determine their effectiveness for face authentication. Face sample quality is a measurement of these intrinsic characteristics. Face sample quality has significant impact on the performance of a face-based biometric system. Recognizing face samples of poor quality is a challenging problem. A number of factors can contribute toward degradation in face sample quality. They include, but not limited to, illumination variation, pose variation, facial expression change, face occlusion, low resolution, and high sensing noise.

Introduction

A typical face-based biometric system operates by capturing face data (images or videos), and comparing the obtained face data against face templates of different individuals in a gallery set. While face templates in the gallery set are normally captured under constrained imaging conditions (e.g., from frontal view, at a short distance from the camera, and under consistent illumination), it is unrealistic to assume controlled acquisition of probe face data. Face data captured under uncontrollable environment usually contains many kinds of defects caused by poor illumination, improper face positioning, and imperfect camera sensors [1]. For instance, when face data is captured in a natural outdoor environment, inconsistent illumination is typically cast on human faces resulting in uneven, extremely strong or weak lightings. Face rotation can also cause significant appearance variations, and at the extreme, face can be self occluded (Fig. 1). When distances between human faces and cameras increase, captured face data will be at low resolution, in low contrast, and likely to contain high imaging noise. In some instances people may wear sunglasses, have varying facial expression, and be with heavy makeup. All of these factors contribute toward potential degradation in the quality of captured face samples, resulting in disparities to those of face templates stored in the gallery set.

Face sample quality has significant impact on the performance of face-based biometric systems.



Face Sample Quality. Figure 1. Face samples of illumination and pose variations from AR and UMIST databases.

Assessing the quality of face samples before applying them in any biometric system may help improve the authentication accuracy. For example, an intruder may wear sunglasses intending to disguise himself, quality assessment of intruder's face samples can give an alert to such a situation. Quantitative measures on the quality of face samples can also be integrated into biometric systems to increase or decrease relevant thresholds. In a people enrollment stage, such quantitative measures of quality also help procure gallery face templates of good quality. Many approaches assess face sample quality using general image properties including contrast, sharpness, and illumination intensity [2]. However, these properties cannot properly measure face sample degradation caused by inconsistent illumination, face rotation, or large face-camera distance. There are a few recent works assessing face sample quality by considering such kinds of degradation. For example in [1], facial-symmetry-based methods are used to measure facial asymmetries caused by nonfrontal lighting and improper facial pose.

When only poor quality face data can be acquired at the authentication stage, face recognition becomes significantly more challenging because of: (1) *Illumination variation* to which the performance of most existing face recognition algorithms and systems is highly sensitive. It has been shown both experimentally [3] and theoretically [4] that face image differences resulting from illumination variation are more significant than either inherent face differences between different individuals, or those from varying face poses [5]. State of the art approaches addressing this problem include heuristic methods, reflectance-model methods, and 3D-model-based methods [6]. Although performance improvement is achieved, none of these methods are truly illumination invariant. (2) *Pose variation* which causes face recognition accuracy to decrease significantly, especially when large pose variations between gallery and probe faces are present. The difficulties would further increase if only an unknown single pose is available for each probe face. In such a situation, an extra independent training set, different from the galley set and containing multiple face images of different individuals under varying poses, will be helpful. Three-dimensional face model or statistical relational learning between different poses can be employed to generate virtual face poses. By generating virtual poses, one can either normalize probe faces of varying poses to a predefined pose, e.g., frontal, or expand the gallery to cover large pose variations. (3) *Low resolution* face data will be acquired when face-camera distances increase, which is rather typical in surveillance imagery. The performance of existing face recognition systems decreases significantly when the resolution of captured face data is reduced below a certain level. This is because the missing high-resolution details in facial appearances and image features make facial analysis and recognition ineffective, either by human operators or by automated systems. It is therefore useful to generate high-resolution face images from low-resolution ones. This technique is known as face hallucination [7] or face ► super-resolution.

Assessment of Face Sample Quality

The performance of face authentication depends heavily on face sample quality. Thus the significance of face sample quality assessment and standardization grows as more practical face-based biometric systems are required. Quality assessment of probe face samples can either reject or accept a probe to improve later face verification or identification accuracy. Quantitative assessment of face sample quality can also be used to assign weights in a biometric fusion scheme.

ISO/IEC WD 29794-1 [8] considers that biometric sample quality can be defined by character (inherent features), fidelity (accuracy of features), or utility (predicted biometrics performance). Many efforts have been made on biometric sample quality assessment for fingerprint, iris, or face data. Most of those on face data are based on general image properties including contrast, sharpness, and illumination intensity [2]. However, the face sample degradation that severely affects face authentication accuracy is from uncontrollable imaging conditions that cause illumination variations, head pose changes, and/or very low-resolution facial appearances. There are a few attempts made on assessing face sample quality caused by these kinds of degradation.

In [9], two different strategies for face sample quality assessment are considered: one is for illumination variation and pose change, another is for facial expression change. In the first strategy, specific measures are defined to correlate with levels of different types of face sample degradation. A polynomial function is then utilized based on each measure for predicting the

performance of a ▶ Eigenface technique on a given face sample. Quality goodness is assessed by selecting a suitable threshold. Since the measurement of facial expression intensity is difficult, in the second strategy, a given face sample is classified into good or poor quality based on its coarse similarity to neutral facial expression. Then the training procedure for each class is achieved by dividing the training set into two subsets, based on whether the samples are recognizable by the Eigenface technique. Then these two subsets are described by Gaussian mixture models (GMMs). In [1], facial-symmetry-based quality scores are used to assess facial asymmetries caused by non-frontal lighting and improper facial pose. In particular, local binary pattern (LBP) histogram features are applied to measure the lighting and pose asymmetries. Moreover, the inter-eye distance is also used to estimate the quality score for whether a face is at a proper distance from the camera.

Recognizing Face Samples of Poor Quality

In general, face recognition under varying illumination is difficult. Although existing efforts to address this challenge have not led to a fully satisfactory solution for illumination invariant face recognition, some performance improvements have been achieved. They can be broadly categorized into: heuristic methods, reflectance-model methods, and 3D-model-based methods [6]. A typical heuristic method applies subspace learning, e.g., principal component analysis (PCA), using training face samples. By discarding a few most significant, e.g., the first three, principal components, variations due to lighting can be reduced. Reflectance-model methods employ a Lambertian reflectance model with a varying albedo field, under the assumption of no attached and cast shadows. The main disadvantage of this approach is the lack of generalization from known objects to unknown objects [10]. For 3D-face model-based approaches, more stringent assumptions are often made and it is also computationally less reliable. For example in [11], it is assumed that the 3D face geometry lies in a linear space spanned by the 3D geometry of training faces and it uses a constant albedo field. Moreover, 3D model-based methods require complex fitting algorithms and high-resolution face images.

There are also attempts to address the problem of face recognition across varying facial poses. In real-world applications, one may have multiple face samples of varying poses in training and gallery sets (since they can be acquired offline), while each captured probe face can only be at an unknown single pose. Three-dimensional model-based methods [12] or statistical learning-based methods can be used to generate virtual face poses [13], by which either probe faces can be normalized to a predefined pose, e.g. frontal view, or gallery faces can be expanded to cover large pose variations. For example in [12], a 3D morphable model is used. The specific 3D face is recovered by simultaneously optimizing the shape, texture, and mapping parameters through an analysis-by-synthesis strategy. The disadvantage of 3D model-based methods is slow speed for real-world applications. Learning-based methods try to learn the relations between different facial poses and how to estimate a virtual pose in 2D domain, e.g., the view-based active appearance model (AAM) [14]. This method depends heavily on the accuracy of face alignment, which unfortunately introduces another open problem in practice.

When the resolution of captured face data falls below a certain level, existing face recognition systems will be significantly affected. Face super-resolution techniques have been proposed to address this challenge. Reconstruction-based approaches require multiple, accurately aligned low-resolution face samples to obtain a high-resolution face image. Their magnification factors of image resolution are however limited [7]. Alternatively, learning-based face super-resolution approaches model high-resolution training faces and learn face-specific prior knowledge from them. They use the learned model prior to constrain the super-resolution process. A super-resolution factor as high as 4×4 can be achieved [7]. The face super-resolution process can also be integrated with face recognition. For example in [15], face image super-resolution is transferred from pixel domain to a lower dimensional eigenface space. Then the obtained high-resolution face features can be directly used in face recognition. Simultaneous face super-resolution and recognition in ▶ tensor space have also been introduced [16]. Given one low-resolution face input of single modality, the proposed method can integrate and realize the tasks of face super-resolution and recognition across different facial modalities including varying facial expression, pose, or illumination. This has been further generalized to unify automatic alignment with super-resolution [17].

Summary

Many face-based biometric systems have been deployed in applications ranging from national border control to building door access, which normally solve the sample quality problem at the initial face acquisition stage. Given ongoing progress on standardization of face sample quality and technical advancement in authenticating face samples of poor quality, the availability of more reliable and convenient face authentication systems is only a matter of time.

Related Entries

▶ Biometric Sample Quality
▶ Face Pose Analysis
▶ Face Recognition

References

1. Gao, X.F., Li, S.Z., Liu, R., Zhang, P.R.: Standardization of face image sample quality. In: Proceedings of Second International Conference on Biometrics (ICB), pp. 242–251. Seoul, Korea (2007)
2. Brauckmann, M., Werner, M.: Technical report. In: Proceedings of NIST Biometric Quality Workshop. (2006)
3. Adini, Y., Moses, Y., Ullman, S.: Face recognition: the problem of compensating for changes in illumination direction. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 721–732 (1997)
4. Zhao, W., Chellappa, R.: Robust Face Recognition Using Symmetric Shape-from-Shading. Technical Report, Center for Automation Research, University of Maryland (1999)
5. Tarr, M.J., Bulthoff, H.H.: Image-based object recognition in man, monkey and machine. Cognition **67**, 1–20 (1998)
6. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: a literature survey. ACM Comput. Surv. **35**(4), 399–458 (2003)
7. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Trans. Pattern Anal. Mach. Intell. **24**(9), 1167–1183 (2002)
8. ISO/IEC JTC 1/SC 37 N 1477: Biometric Sample Quality Standard – Part 1: Framework (2006)
9. Abdel-Mottaleb, M., Mahoor, M.H.: Application notes - algorithms for assessing the quality of facial images. IEEE Comput. Intell. Mag. **2**(2), 10–17 (2007)
10. Baker, S., Kanade, T.: Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **29**(2), 230–245 (2007)
11. Atick, J., Griffin, P., Redlich, A.: Statistical approach to shape from shading: reconstrunction of 3-dimensional face surfaces from single 2-dimentional images. Neural Comput. **8**(2), 1321–1340 (1996)
12. Blanz, V., Vetter, T.: Face recognition based on fitting a 3-D morphable model. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1063–1074 (2003)
13. Li, Y., Gong, S., Liddell, H.: Constructing facial identity surfaces for recognition. Int. J. Comput. Vision **53**(1), 71–92 (2003)
14. Cootes, T.F., Walker, K., Taylor, C.J.: View-based active appearance models. In: Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition (FG), pp. 227–232. Grenoble, France (2000)
15. Gunturk, B.K., Batur, A.U., Altunbasak, Y., Hayes, M.H., Mersereau, R.M.: Eigenface-Domain Super-Resolution for Face Recognition. IEEE Transactions on Image Processing, **12**(5), 597–606 (2003)
16. Jia, K., Gong, S.: Multi-modal tensor face for simultaneous super-resolution and recognition. In: Proceedings of Tenth International Conference on Computer Vision (ICCV), pp. 1683–1690. Beijing, China **2** (2005)
17. Jia, K., Gong, S.: Generalised face super-resolution. IEEE Transactions on Image Processing, **17**(6), 873–886(2008).

Face Sample Standardization

▶ Face Sample Quality

Face Sample Synthesis

Sami Romdhani, Jasenko Zivanov
Computer Science Department, University of Basel, Basel, Switzerland

Synonyms

Face image synthesis; Rendering; Image formation process

Definition

Face Sample Synthesis denotes the process of generating the image of a human face by a computer program. The input of this process is a set of parameters that

describes (1) the position from which the face is viewed, (2) the illumination environment around the face, (3) the identity of the person, and (4) the expression of the person. Other parameters may also be used such as the age of the person, parameters describing the make-up, etc. The output is an image of a human face.

Introduction

Face Sample Synthesis denotes the process of generating the image of a human face by a computer program. Optimally, this image should be realistic and virtually indistinguishable from a photography of a live scene. Additionally, the computer program should be generic: able to synthesize the face of any individual, viewed from any pose and illuminated by any arbitrarily complex environment. The objective of this article is to review the techniques used to reach this goal. It may also be desirable to generate faces with different expressions, different attributes such as makeup style or facial hair. One might also want to render image sequences with realistic facial motion. However, it is outside the scope of this article to address the methods enabling such synthesizes.

A photograph of a face is a projection onto an image plane of a 3D object. The intensity of a pixel of this photograph directly depends on the amount of light that is reflected from the object point imaged at the pixel location. Thus, this article first reviews 3D to 2D projections (the finite projective camera model) and illumination modeling (the Lambertian and Phong light reflection models usually used in face recognition systems). Then, the basics of identity modeling are summarized. At the end of the article, the reader will have an overview of the process required to synthesize a face image from any individual, viewed from any angle, and illuminated from any direction.

Research on computer-based face recognition dates back from the 1970s. In those times, most popular methods (e.g., [1]) were based on distances and angles between landmark points (such as eyes and mouth corners, nostril, chin top, etc.). Then, in the beginning of the 1990s, the appearance-based methods came in and quickly attracted most of the attention [2]. Contrasting with the former landmark points methods, these techniques use the entire face area for recognition. They are based on a prior generative model capable of synthesizing a face image given a small number of parameters. Analysis is performed by estimating the parameters, denoted by $\hat{\theta}$, which synthesize a face image that is as similar as possible to the input image. Hence, these methods are called *Analysis by Synthesis*. This is usually done using a sum of square error functions:

$$\hat{\theta} = \arg\min_{\theta} \sum_{i} \|I_{i,input} - I_{i,model}(\theta)\|^2, \qquad (1)$$

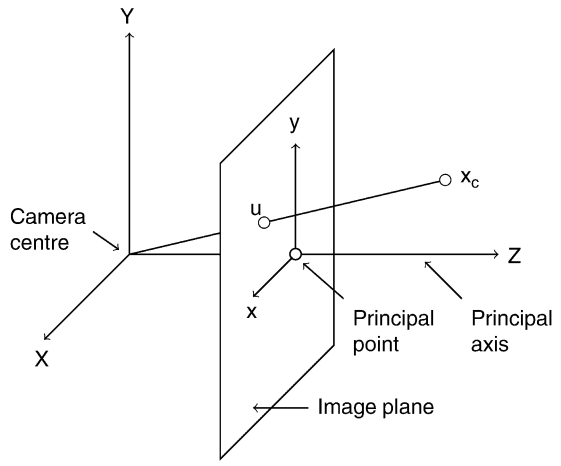
where the index i represent pixel i and the sum runs over all pixels of the face area. The formation of the model image, $I_{model}(\theta)$, is the topic of this article. Initially, the models used a 2D representation of the face structure [2], however, in order to account for pose and illumination variation, it is accepted that 3D models provide more accurate results [3, 4]. Hence, this article reviews the process of generating a face image from a 3D model.

Four ingredients are necessary to synthesize a face from a 3D representation [5]: The face surface of the individual to be imaged must be sampled across a series of points resulting in a list of 3D vertices. Obtaining a surface from a list of vertices is achieved by a triangle list that connects triplets of vertices. The triangle list defines the topology of the face. It is used, among other things, to compute surface normals and the visibility of a surface points using (for instance) a "Z-buffer" visibility test.

The third constituent is the color of the face. It can be represented by an RGB color for a dense set of surface points. These surface points are called "texels." If the texels are the same points as the vertices, then the color model is called "per vertex color." Alternatively, a much denser texel sampling can be used and the texels are arranged in a "texture map." In order to synthesize unconstrained illumination images, the texels must be free of any illumination effect and code the "albedo" of a point. The albedo is defined as the diffuse color reflected by a surface point. Finally, the last ingredient is a reflection model that relates the camera direction and the intensity of light reflected by a surface point, to the intensity, the direction, and the wavelength of light reaching the point.

Finite Projective Camera Model

This section briefly describes how a 3D object is imaged on a 2D image.



In Computer Vision and Computer Graphics, a finite projective camera model is usually chosen. This camera follows a central projection of points in space onto a plane. For now, lets assume that the camera is at the origin of an Euclidean coordinate system and that it is pointing down the Z-axis. In that system, a 3D point, $\mathbf{x}_c = (X_c, Y_c, Z_c)^t$ is projected onto a 2D point, $\mathbf{u}$, in the image frame according to the following equation, in which the focal length, denoted by f, is the distance between the camera center and the principal point: $\mathbf{u} = (fX_c/Z_c, fY_c/Z_c)$. This equation assumes that the origin of the image plane coordinate system is at the principal point. In general, it might not be and, denoting the coordinates of the principal point by (p_x, p_y), the mapping becomes:

$$\mathbf{u} = (fX_c/Z_c + p_x, fY_c/Z_c + p_y). \tag{2}$$

In a face image synthesis, the point $\mathbf{x}_c$ is one vertex of the 3D shape of a face in *camera coordinate frame*. It is easier to represent the ensemble of vertices of the face in an *object coordinate frame*. The origin of this frame is attached to the object, a typical choice is to locate it at the center of mass of the face. The 3D coordinate of the camera center in the object frame is denoted by $\mathbf{c}$.

Additionally, the object coordinate frame is generally not aligned with the camera coordinate frame, i.e., the face is not always frontal. The rotation between the face and the camera is denoted by the 3×3 matrix $\mathbf{R}$. It can be represented by a product of rotations along the coordinate axes of the object frame:

$$\mathbf{R}_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{pmatrix},$$
$$\mathbf{R}_\beta = \begin{pmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{pmatrix},$$
$$\mathbf{R}_\gamma = \begin{pmatrix} \cos(\gamma) & \sin(\gamma) & 0 \\ -\sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{R} = \mathbf{R}_\alpha \mathbf{R}_\beta \mathbf{R}_\gamma. \tag{3}$$

The relation between the object and camera frames is then: $\mathbf{x}_c = \mathbf{R}(\mathbf{x} - \mathbf{c})$. It is often convenient not tomake the camera center explicit and to introduce $\mathbf{t} = -\mathbf{Rc}$. In this case, the relation is simply:

$$\mathbf{x}_c = \mathbf{Rx} + \mathbf{t}. \tag{4}$$

As a result, projecting a point $\mathbf{x}$ in object coordinate frame onto the image plane is summarized by the following expression, in which $\mathbf{R}_i$ denotes the row number i of the matrix $\mathbf{R}$.

$$\begin{cases} u_x = f\frac{\mathbf{R}_1\mathbf{x}+\mathbf{t}_x}{\mathbf{R}_3\mathbf{x}+\mathbf{t}_z} + p_x \\ u_y = f\frac{\mathbf{R}_2\mathbf{x}+\mathbf{t}_y}{\mathbf{R}_3\mathbf{x}+\mathbf{t}_z} + p_y \end{cases} \tag{5}$$

Estimating the parameters of a finite projective camera model requires then the estimation of nine parameters: f, α, β, γ, $\mathbf{t}_x$, $\mathbf{t}_y$, $\mathbf{t}_z$, p_x, p_y. Note that in this explanation some subtle parameters that have only a minor effect on the synthesis and on the analysis by synthesis results are neglected: Some CCD cameras do not have square pixels (two additional parameters) and the skew parameter that is zero for most normal cameras [6].

Lighting Model

The previous section showed where, in the image, to draw a surface point from its 3D coordinates. Now the question is: What pixel value to draw on this point? The pixel value is the intensity of the light reflected by the surface point, which is computed using a lighting model. Much of the realism of a rendering depends on the ▶ lighting model.

This model, in turn, depends on three factors: The number and type of light sources, the reflectance function, and the method used to compute surface normals. Light modeling is still undergoing considerable research efforts in the computer graphics community (the main challenge being to make photo-realistic rendering algorithms computationally efficient). In this article, the fundamental notions are only briefly introduced.

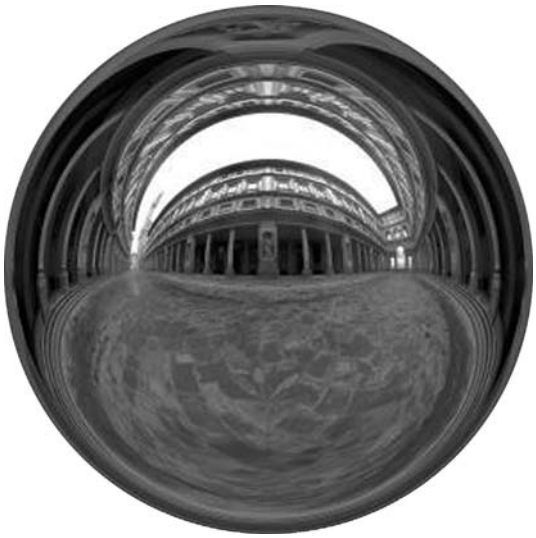
If light is emitted from direction $\bar{l}$ with intensity l, then the quantity of light received by an infinitesimally small surface patch around surface point $\mathbf{x}$ is $\langle \bar{n}_x, \bar{l} \rangle \; l$, where $\bar{n}_x$ is the normal of the surface patch at the point $\mathbf{x}$ and $\langle \cdot, \cdot \rangle$ is the scalar product (if it is positive and null otherwise). If the surface point projects onto pixel i of the image, yields

$$I_{i,\text{model}} = r_x(\bar{v}, \bar{l}) \cdot \langle \bar{n}_x, \bar{l} \rangle \cdot l \cdot S_{x,\bar{l}}, \tag{6}$$

where $r_x(\cdot)$ denotes the reflectance function at point x and $\bar{v}$, the viewing direction (defined as the direction from the point to the camera center). $S_{x,\bar{l}}$ denotes the cast shadow binary variable: If there is another object or if some part of the face is between point $\mathbf{x}$ and the point at infinity in direction $\bar{l}$, then the light is shadowed at the point, and $S_{x,\bar{l}}$ is zero, otherwise it is equal to one. Cast shadows are usually computed by a shadow map [7].

Light Source

The simplest and most computationally efficient is to use one directed light source at infinity and one ambient light source. The light reflected by a surface point from an ambient light source does not depend on the local surface around the point, it only depends on the albedo of the point. In real world, however, a perfectly ambient light never exists and it rarely happens that a point is illuminated only by a single light source. Indeed, light emanating from a light source might bounce off a wall, for instance, and then reach the object point. Hence, in real world, light comes from all directions. An environment map [8] is usually used to model this effect. It codes the intensity of light reaching an object for a dense sampling of directions. It is acquired by photographing a mirrored sphere [9]. Due to the additive nature of light, rendering with several light sources (as is the case for environment maps) is performed by summing (or integrating) over the light sources:



$$I_{i,\text{model}} = \sum_j r_x(\bar{v}, \bar{l}_j) \cdot \langle \bar{n}_x, \bar{l}_j \rangle \cdot l_j \cdot S_{x,\bar{l}_j}. \tag{7}$$

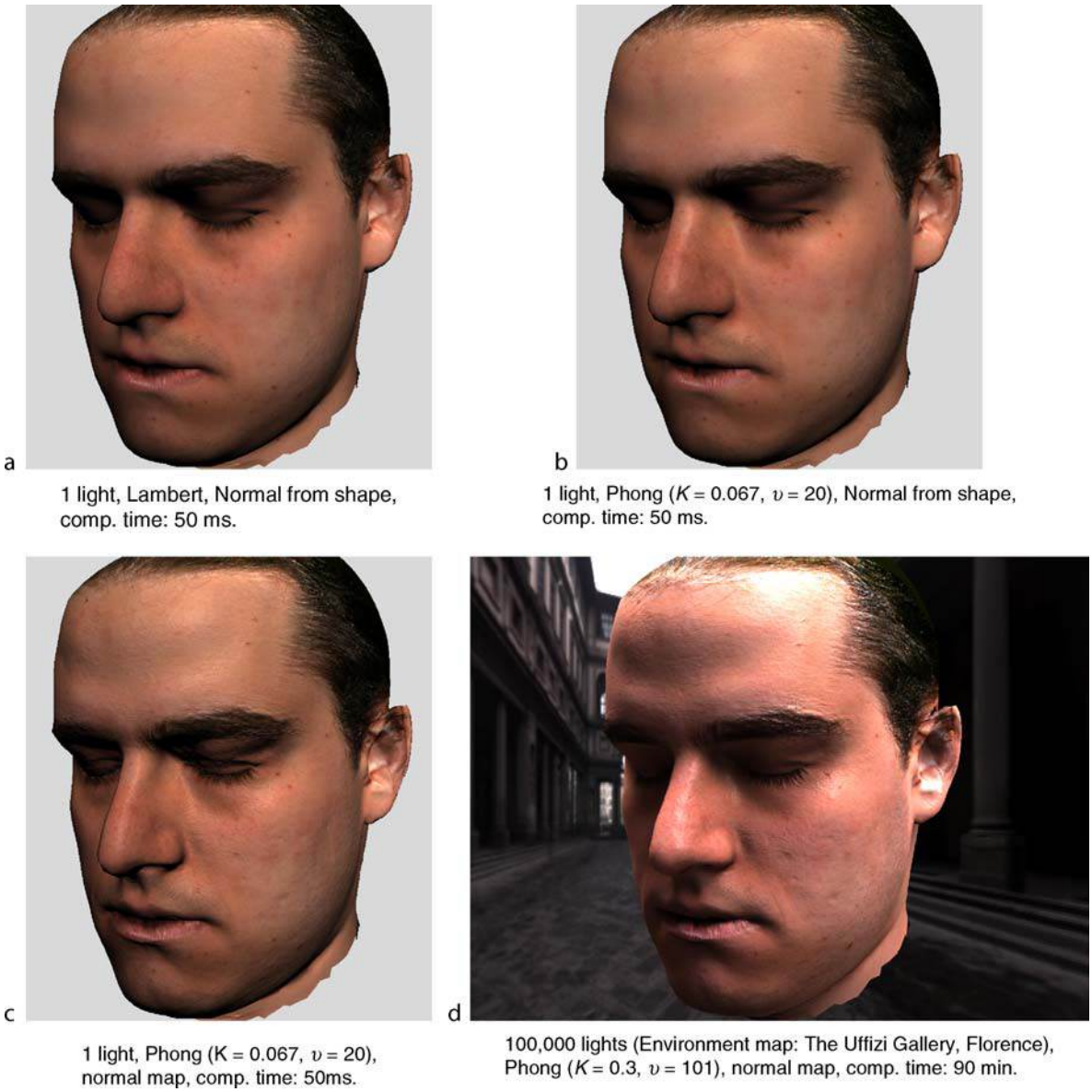
The following image is an environment map acquired at the Uffizi Gallery in Florence, Italy. Each pixel of this photograph is attached to a direction and represent a light source. This environment map is used to illuminate Panel d of Fig. 1.

Reflectance Function

In (6), the four-dimensional reflectance function $r_x(\bar{v}, \bar{l})$ is called the Bidirectional Reflectance Distribution Function (BRDF). It also depends on the wavelength of the incoming light (usually represented by its RGB color). The BRDF describes the properties of the material at point $\mathbf{x}$.

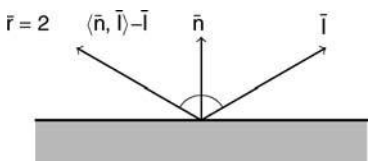
The simplest model of reflectance function is certainly the Lambertian model for which the function is equal to a constant (the albedo at point $\mathbf{x}$). This means that incident light is scattered equally in all directions, which only happens for perfectly diffuse objects (totally matte, without shininess). For human face, this is the case only when the skin is covered by a very fine layer of powder. An example of rendering with a Lambertian reflectance is displayed on Fig. 1a.

Specular reflection takes place when light is reflected at a point without absorption by the material. For perfectly specular material, such as mirrors, light is reflected in only one direction (the reflectance function is a Dirac function): when the viewing angle is equal to the angle



Face Sample Synthesis. **Figure 1** Renderings using different illumination models. The 3D model includes 24,367 vertices and 48,660 triangles.

of incidence . Generally, in Computer Graphics, the reflectance used is a combination of diffuse and specular reflectance. The most well-known model is the Phong model:



$$I^{\text{phong}}_{i,\text{model}} = (\mathbf{c} \cdot \langle \bar{n}_x, \bar{l} \rangle + K \cdot \langle \bar{v}, r \rangle^{\bar{\nu}}) \cdot l \cdot S_{x,\bar{l}}$$

where $\mathbf{c}$ is the albedo at point $\mathbf{x}$, $\bar{r}$ is the reflection direction (depending on the normal and the lighting direction, as shown on the sketch), K is the fraction of energy specularly reflected, and ν is an index that controls the "tightness" of the specular highlight (note that there is one such equation for each color channel). In this equation, the first summand inside the brackets is the diffuse (i.e. Lambertian) part and the second

one is the specular part. Modeling the specular highlight is important for human skin, as it often has a thin layer of oil or sweat above the pigmented cells. Figure 1a–d show examples of face rendering with a Phong model.

The Pong model assumes smooth surfaces, but, in reality, surfaces are imperfect and exhibit microgeometry. There exist more complex BRDF models that represent the surface as composed of micro-facets that can shadow and mask each other. Another effect is the off-specular highlight: When the angle of incidence is grazing (near 90°), some materials (such as human skin) reflect much more light than is absorbed, causing the color of the point to approach that of the light. This is accounted for by the Fresnel term that is the ratio between reflected and absorbed light. Some of the more complex BRDF model that accounts for these two effects are Blinn [10], Cook-Torrance [11], Torrance-Sparrow [12], and more recently Lafortune [13].

Note that this is not the end of the story, yet. The BRDF assumes that the outgoing light at one point results only from the incoming light at the same point. This is in fact an approximation as it neglects the scattering of light within the material. This phenomenon is modeled by the bidirectional surface scattering reflectance distribution function (BSSRDF) [14] of which the BRDF is a special case. Human skin does show some important subsurface scattering effects and to reach photo-realism these effects should not be neglected. This is for instance apparent when the ear is illuminated from the back. It then looks translucent which results from the subsurface scattering.

Normals

Human skin is not a smooth surface. Pores and wrinkles induce very small scale variations of the surface. Representing these variations with 3D vertices would require a very fine sampling of the head resulting in an overly large number of vertices making the rotation or visibility test computationally inefficient. The concept of "normal mapping" was developed precisely for this reason. Instead of computing a normal from the shape (by interpolating the normals of the triangles corners in which a pixel is projected from), the normals are computed from a dense normal map (for which interpolation is carried similarly to the texture map). A normal map is generally acquired by photometric stereo [15]: Several photographs of the face of a subject are taken with different light directions. The subject and the camera must be perfectly still during this acquisition process (a pixel must be the projection of exactly the same point on the subject face for all photographs). Each photograph yields one measurement of the BRDF of a surface point. Using several measurements a BRDF model can be fitted, thereby recovering the normal of the point. Often, for its simplicity, a Lambertian model is used, in which case, the operator must choose the light direction such as to minimize specular reflections.

Figure 1 shows different types of rendering from the most simple (left) to more complex and realistic (right).

Identity Modeling

So far, an overview of the face image synthesis process from a 3D model of the face of an individual is presented. This 3D model can be acquired by a 3D scanner or can be manually crafted with a modeling software. These processes can be tedious and expensive. Therefore, it is desirable to be able to *generate* the 3D shape and texture (i.e., albedo) from *any individual*. This can be done by defining a vector space of shapes and of textures and probability distributions in these spaces. This is accomplished by learning typical face variations from an example set of 3D faces. The vector space is defined by densely registering the examples with a reference face, thereby defining a label for each vertex. Once the vector spaces are defined, linear combination of the example shapes and textures are made to generate the shape and texture of new (i.e., out of the example set) individuals. The coefficients of these linear combinations are the parameters of the identity model. One individual is coded by a specific value for each parameter. However, some variations are more typical than others and probability distributions in the vector space must be used in order to ensure the plausibility of a novel individual. If the probability distributions of the human faces in the vector spaces are assumed Gaussian, then the most efficient coding is yielded by a Principal Component Analysis [16] of the examples. These principles are used by the *3D Morphable Model* [3], the state of the art identity generic human face model.

Denoting by $\mathbf{X}$ the $3 \times N$ matrix with the 3D position of N vertices (hence the position of a vertex

x in the "Finite Projective Camera Model" section is a column of the matrix **X**) and by **C** the $3 \times N$ matrix with the RGB albedo of N vertices, a novel 3D face is yielded by the following equations:

$$\mathbf{X} = \sum_{i}^{M} \alpha_i \mathbf{X}_i, \qquad \mathbf{C} = \sum_{i}^{M} \beta_i \mathbf{C}_i, \tag{8}$$

where $\mathbf{X}_i$ and $\mathbf{C}_i$ are the M shape and texture principal components and α_i and β_i the shape and texture parameters.

Conclusion

In conclusion, the set of parameters θ of an analysis by synthesis method (1) is composed of nine parameters for the projection. For illumination, using a Phong reflectance model with one light source and with normals computed from the shape, seven parameters must be estimated: three parameters for the intensity of the colored light, two parameters for its direction along with the specular coefficient K and the Phong exponent ν. Additionally, $2M$ parameters must be recovered for the 3D shape and the texture.

Using a normal map model (generic for all individuals) and an environment map for analysis by synthesis has never been attempted. Indeed, it would result in a tremendously complicated problem for the following reasons: It is unclear how to model normal maps that would generalize for any individual. A simple linear combination as is used for the shape and texture cannot be used for normals because a normal vector is a unit length vector and the sum or the mean of two unit length vectors does not result in a unit length vector. Moreover, defining correspondences for pores and wrinkles (which would be required to make a vector space and avoid blur results) is for the moment unsolved. As for the light sources, estimating the direction and intensity of a single light source from a single facial image is already an ill-posed problem (there is not enough information in one image to completely constraint the solution), let alone with a large number of light sources as is the case when using an environment map.

Summary

The motivation of face sample synthesis is not only to generate face images from a small number of parameters but also to analyze them using an analysis by synthesis approach. In this article, the two main sources of face image variations (pose and illumination) are accounted for by a finite projective camera model. Illumination modeling is more complicated and requires the operator to choose the type of illumination sources, the type of reflectance function, and the manner to generate normals (either from the shape or acquired by a photometric stereo method). Finally, identity variations can be obtained by a linear combination of examples.

Related Entries

- ▶ Deformable Models
- ▶ Face Pose Analysis

References

1. Kanade, T.: Computer Recognition of Human Faces. Birkhäuser Verlag, Stuttgart, Germany (1973)
2. Cootes, T., Edwards, G., Taylor, C.: Active appearance model. Fifth European Conference on Computer Vision. Freiburg, Germany (1998)
3. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D-faces. SIGGRAPH 99. Los Angeles, California, USA (1999)
4. Romdhani, S., Vetter, T.: Estimating 3D shape and texture using pixel intensity, edges, specular highlights, texture constraints and a prior. CVPR. San Diego, CA, USA (2005)
5. Parke, F.I., Waters, K.: Computer Facial Animation. AKPeters Wellesley, Massachusetts, USA (1996)
6. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2000)
7. Woo, A., Poulin, P., Fournier, A.: A survey of shadow algorithms. IEEE Comput. Graph. Appl. **10**(6), 13–32 (1990)
8. Greene, N.: Environment mapping and other applications of world projections. IEEE Comput. Graph. Appl. **6**(11) (1986)
9. Debevec, P., Malik, J.: Recovering high dynamic range radiance maps from photographs. Siggraph. Los Angeles, California, USA (1997)
10. Blinn, J.F.: Models of light reflection for computer synthesized pictures. SIGGRAPH '77: Proceedings of the Fourth Annual Conference on Computer Graphics and Interactive Techniques, pp. 192–198. New York, NY, USA, ACM (1977)
11. Cook, R.L., Torrance, K.E.: A reflectance model for computer graphics. ACM Trans. Graph. **1**(1), 7–24 (1982)
12. Torrance, K.E., Sparrow, E.M.: Theory for off-specular reflection from roughened surfaces, pp. 32–41. Radiometry (1992)
13. Lafortune, E.P.F., Foo, S.C., Torrance, K.E., Greenberg, D.P.: Non-linear approximation of reflectance functions. SIGGRAPH '97: Proceedings of the 24th Annual Conference on Computer

Graphics and Interactive Techniques, pp. 117–126. New York, NY, USA, ACM /Addison-Wesley (1997)
14. Jensen, H.W., Marschner, S.R., Levoy, M., Hanrahan, P.: A practical model for subsurface light transport. SIGGRAPH, ACM/ Addison-Wesley (2001)
15. Barsky, S., Petrou, M.: Colour photometric stereo: simultaneous reconstruction of local gradient and colour of rough textured surfaces. ICCV p. 600 (2001)
16. Jolliffe, I.T.: Principal Component Analysis. Springer, Berlin (2002)

Face Sample Utility

► Face Sample Quality

Face Sketching

A face sketching is a parsimonious yet expressive representation of face. It depicts concise sketches of face that captures the most essential perceptual information with a number of strokes.

► And-Or Graph Model for Faces

Face Tracking

Amit K. Roy-Chowdhury, Yilei Xu
Department of Electrical Engineering, University of California, Riverside, CA, USA

Synonym

Facial motion estimation

Definition

In many face recognition systems, the input is a video sequence consisting of one or more faces. It is necessary to track each face over this video sequence so as to extract the information that will be processed by the recognition system. Tracking is also necessary for 3D model-based recognition systems, where the 3D model is estimated from the input video. Face tracking can be divided along different lines depending upon the method used, e.g., head tracking, feature tracking, image-based tracking, model-based tracking. The output of the face tracker can be the 2D position of the face in each image of the video (2D tracking), the 3D pose of the face (3D tracking), or the location of features on the face. Some trackers are also able to output other parameters related to lighting or expression. The major challenges encountered by face tracking systems are robustness to pose changes, lighting variations, and facial deformations due to changes of expression, occlusions of the face to be tracked and clutter in the scene that makes it difficult to distinguish the face from the other objects.

Introduction

Tracking, which is essentially ► motion estimation, is an integral part of most face processing systems. If the input to a face recognition system is a video sequence, as obtained from a surveillance camera, tracking is needed to obtain correspondence between the observed faces in the different frames and to align the faces. It is so integral to video-based face recognition systems that some existing methods integrate tracking and recognition [1]. It is also a necessary step for building 3D face models. In fact, tracking and 3D modeling are often treated as two parts of one single problem [2–4].

There are different ways to classify face tracking algorithms [5]. One such classification is based on whether the entire face is tracked as a single entity (sometimes referred to as head tracking) or whether individual facial features are tracked. Sometimes a combination of both is used. Another method of classification is based on whether the tracking is in the 2D image space or in 3D pose space. For the former, the output (overall head location or facial feature location) is a region in the 2D image and does not contain information about the change in the 3D orientation of the head. Such methods are usually not very robust to changes of pose, but are easier to handle computationally. Alternatively, 3D tracking methods, which work by fitting a 3D model to each image of the video, can provide estimates of the 3D pose of the face. However, they are

usually more computationally intensive. Besides, many advanced face tracking methods are able to handle challenging situations like facial ▶ deformations, changes of lighting, and partial occlusions.

A broad overview of the basic mathematical framework of face tracking methods will be given first, followed by a review of the current state-of-the-art and technical challenges. Next, a few application scenarios will be considered, like surveillance, face recognition, and face modeling, including discussion of the importance of face tracking in each of them. Then some examples of face tracking in challenging situations will be shown, before conclusion.

Basic Mathematical Framework

An overview of the basic mathematical framework that explains the process in which most trackers work is provided here. Let $\mathbf{p} \in \Re^p$ denote a parameter vector, which is the desired output of the tracker. It could a 2D location of the face in the image, the 3D pose of the face, or a more complex set of quantities that also include lighting and deformation parameters. Define a synthesis function $f\colon \Re^2 \times \Re^p \rightarrow \Re^2$ that can take an image pixel $\mathbf{v} \in \Re^2$ at time $(t-1)$ and transform it to $f(\mathbf{v}, \mathbf{p})$ at time t. For a 2D tracker, this function f could be a transformation between two images at two consecutive time instants. For a 3D model-based tracker, this can be considered as a rendering function of the object at pose $\mathbf{p}$ in the camera frame to the pixel coordinates $\mathbf{v}$ in the image plane. Given an input image $I(\mathbf{v})$, align the synthesized image with it so as to obtain

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} g(f(\mathbf{v}, \mathbf{p}) - I(\mathbf{v})), \qquad (1)$$

where $\hat{\mathbf{p}}$ denotes the estimated parameter vector for this input image $I(\mathbf{v})$.

The essence of this approach is the well-known Lucas–Kanade tracking, an efficient and accurate implementation of which has been proposed using the inverse compositional approach [6]. Depending on the choice of $\mathbf{v}$ and $\mathbf{p}$, the method is applicable to the overall face image, a collection of discrete features, or a 3D face model. The ▶ cost function g is often implemented as an L_2 norm, i.e., the sum of the squares of the errors over the entire region of interest. However, other distance metrics may be used. Thus a face tracker is often implemented as a least-squares ▶ optimization problem.

Let us consider the problem of estimating the change, $\triangle \mathbf{p}_t \triangleq \mathbf{m}_t$, in the parameter vector between two consecutive frames, $I_t(\mathbf{v})$ and $I_{t-1}(\mathbf{v})$ as

$$\hat{\mathbf{m}}_t = \arg\min_{\mathbf{m}} \sum_{\mathbf{v}} (f(\mathbf{v}, \hat{\mathbf{p}}_{t-1} + \mathbf{m}) - I_t(\mathbf{v}))^2, \qquad (2)$$

and

$$\hat{\mathbf{p}}_t = \hat{\mathbf{p}}_{t-1} + \hat{\mathbf{m}}_t. \qquad (3)$$

The optimization of the above equation can be achieved by assuming a current estimate of $\mathbf{m}$ as known and iteratively solve for increments $\triangle \mathbf{m}$ such that

$$\sum_{\mathbf{v}} (f(\mathbf{v}, \hat{\mathbf{p}}_{t-1} + \mathbf{m} + \triangle \mathbf{m}) - I_t(\mathbf{v}))^2 \qquad (4)$$

is minimized.

Performance Analysis

While the basic idea of the face tracking algorithms is simple, the challenge comes in being able to perform the optimization efficiently and accurately. The function, f, will be nonlinear, in general. This is because f will include camera projection, the 3D pose of the object, the effect of lighting, the surface reflectance, nonrigid deformations, and other factors. For example, in [7], the authors derived a bilinear form for this function under the assumption of small motion. It could be significantly more complex in general. This complexity makes it difficult to obtain a global optimum for the optimization function, unless a good starting point is available. This initialization is often obtained through a face detection module working on the first frame of the video sequence. For 3D model-based tracking algorithms, it also requires registration of the 3D model to the detected face in the first frame.

The need for a good initialization for stable face tracking is only one of the problems. All trackers suffer from the problem of drift of the estimates and face tracking is no exception. Besides, the synthesis function f may be difficult to define precisely in many instances. Examples include partial occlusion of the face, deformations due to expression changes, and variations of lighting including cast shadows. Special care needs to be taken to handle these situations, since direct optimization of the cost function (2) would give an incorrect result.

Computational speed is another important issue in the design of tracking algorithms. Local optimization methods like gradient descent, Gauss–Newton, and Levenberg–Marquardt [8] can give a good result if the starting point is close to the desired solution. However, the process is often slow because it requires recomputation of derivatives at each iteration. Recently, an efficient and accurate method of performing the optimization has been proposed by using an inverse compositional approach, which does not require recomputation of the gradients at each step [6]. In this approach, the transformation between two frames is represented by a ▶ Face Warping function, which is updated by first inverting the incremental warp and then composing it with the current estimate. Our independent experimental evaluation has shown that on real-life facial video sequences, the inverse compositional approach leads to a speed-up by at least one order of magnitude, and often more, leading to almost real-time performance in most practical situations.

Challenges in Face Tracking

As mentioned earlier, the main challenges that face tracking methods have to overcome are (1) variations of pose and lighting, (2) facial deformations, (3) occlusion and clutter, and (4) facial resolution. These are the areas where future research in face tracking should concentrate. Some of the methods proposed to address these problems will be reviewed briefly below.

1. *Robustness to pose and illumination variations.* Pose and ▶ illumination variations often lead to loss of track. One of the well-known methods for dealing with illumination variations was presented in [9], where the authors proposed using a parameterized function to describe the movement of the image points, taking into account illumination variation by modifying the brightness constancy constraint of optical flow. Illumination invariant 3D tracking was considered within the active appearance model (AAM) framework in [10], but the method requires training images to build the model and the result depends on the quality and variety of such data. 3D model based motion estimation algorithms are the usually robust to pose variations, but often lack robustness to illumination. In [7], the authors proposed a model-based face tracking method that was robust to both pose and lighting changes. This was achieved through an analytically derived model for describing the appearance of a face in terms of its pose, the incident lighting, shape, and surface reflectance. Figure 1 shows an example.
2. *Tracking through facial deformations.* Tracking faces through changes of expressions, i.e., through facial deformations, is another challenging problem. An example of face tracking through changes of expression and pose is shown in Fig. 2. A survey of work on facial expression analysis can be found in [12]. The problem is closely related to modeling of facial expressions, which has applications beyond tracking, notably in computer animation. A well-known work in this area is [13], which has been used by many researchers for tracking, recognition, and reconstruction. In contrast to this model-based approach, the authors in [14] proposed a data-driven approach for tracking and recognition of non-rigid facial motion. More recently, the 3D morphable model [15] has been quite popular in synthesizing different facial expressions, which implies that it can also be used for tracking by posing the problem as estimation of the synthesis parameters (coefficients of a set of basis functions representing the morphable model).
3. *Occlusion and clutter.* As with most tracking problems, occlusion and clutter affect the performance

Face Tracking. Figure 1 Tracked points on a face through changes of pose and illumination. These points are projections of a 3D face mesh model.

Face Tracking. **Figure 2** An example of face tracking under changes of pose and expressions. The estimated pose is shown on the top of the frames. The pose is represented as an unit vector for the rotation axis, and the rotation angle in degrees, where the reference is taken to be the frontal face.

Face Tracking. **Figure 3** Tracked points on a face through changes of scale and illumination.

of most face trackers. One of the robust tracking approaches in this scenario is the use of particle filters [16], which can recover from a loss of track given a high enough number of particles and observations. However, in practice, occlusion and clutter remain serious impediments in the design of highly robust face tracking systems.

4. *Facial resolution.* Low resolution will hamper performance of any tracking algorithm, with face tracking being no exception. In fact, [5] identified low resolution to be one of the main impediments in video-based face recognition. Figure 3 shows an example of tracking through scale changes and illumination. Super-resolution approaches can be used to overcome these problems to some extent. However, super-resolution of faces is a challenging problem by itself because of detailed facial features that need to be modeled accurately. Recently, [17] proposed a method for face super-resolution using AAMs. Super-resolution requires registration of multiple images, followed by interpolation. Usually, these two stages are treated separately, i.e., registration is obtained through a tracking procedure followed by super-resolution. In a recent paper

[18], the authors proposed feeding back the super-resolved texture in the nth frame for tracking the $(n+1)$th frame. This improves the tracking, which, in turn, improves the super-resolution output. This could be an interesting area of future work taking into consideration issues of stability and convergence.

Some Applications of Face Tracking

Some applications where face tracking is an important tool have been highlighted below:

1. *Video surveillance.* Since faces are often the most easily recognizable signature of identity and intent from a distance, video surveillance systems often focus on the face [5]. This requires tracking the face over multiple frames.
2. *Biometrics.* Video-based face recognition systems require alignment of the faces before they can be compared. This alignment compensates for changes of pose. Face tracking, especially 3D pose estimation, is therefore an important component of such applications. Also, integration of identity over the entire video sequence requires tracking the face [1].
3. *Face modeling.* Reconstruction of the 3D model of a face from a video sequence using structure from motion requires tracking. This is because the depth estimates are related nonlinearly to the 3D motion of the object. This is a difficult nonlinear estimation problem and many papers can be found that focus primarily on this, some examples being [2–4].
4. *Video communications and multimedia systems.* Face tracking is also important for applications like video communications. Motion estimates remove the interframe redundancy in video compression schemes like MPEG and H.26x. In multimedia systems like sports videos, face tracking can be used in conjunction with recognition or reconstruction modules, or for focusing on a region of interest in the image.

Summary

Face tracking is an important criterion for a number of applications, like video surveillance, biometrics, video communications, and so on. A number of methods have been proposed that work reasonably well under moderate changes of pose, lighting and scale. The output of these methods vary from head location in the image frame to tracked facial features to 3D pose estimation. The main challenge that future research should address is robustness to changing environmental conditions, facial expressions, occlusions, clutter, and resolution.

Related Entries

- ▶ Face Alignment
- ▶ Face Recognition

References

1. Zhou, S., Krueger, V., Chellappa, R.: Probabilistic recognition of human faces from video. Comput. Vision Image Understand. **91**, 214–245 (2003)
2. Fua, P.: Regularized bundle-adjustment to model heads from image sequences without calibration data. Int. J. Comput. Vision **38**, 153–171 (2000)
3. Shan, Y., Liu, Z., Zhang, Z.: Model-based bundle adjustment with application to face modeling. In: Proceedings of IEEE International Conference on Computer Vision, pp. 644–651 (2001)
4. Roy-Chowdhury, A., Chellappa, R., Gupta, R.: 3D face modeling from monocular video sequences. In: Face Processing: Advanced Modeling and Methods. Academic Press, New York (2005)
5. Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A.: Face Recognition: A Literature Survey. ACM Transactions (2003)
6. Baker, S., Matthews, I.: Lucas–Kanade 20 years on: A unifying framework. Int. J. Comput. Vision **56**, 221–255 (2004)
7. Xu, Y., Roy-Chowdhury, A.: Integrating motion, illumination and structure in video sequences, with applications in illumination-invariant tracking. IEEE Trans. Pattern Anal. Machine Intell. Vol. 29, 793–806 (2007)
8. Luenburger, D.: Optimization by Vector Space Methods. Wiley, New York (1969)
9. Hager, G.D., Belhumeur, P.: Efficient region tracking with parametric models of geometry and illumination. IEEE Trans. Pattern Anal. Mach. Intell. **20**, 1025–1039 (1998)
10. Koterba, S., Baker, S., Matthews, I., Hu, C., Xiao, H., Cohn, J., Kanade, T.: Multi-view aam fitting and camera calibration. In: IEEE Intl. Conf. Comput. Vision (2005)
11. Lepetit, V., Fua, P.: Monocular Model-Based 3D Tracking of Rigid Objects. Now Publishers Inc. (2005)
12. Fasel, B., Luettin, J.: Automatic facial expression analysis: a survey. Pattern Recognit. **86**, 259–275 (2003)
13. Terzopoulos, D., Waters, K.: Analysis and synthesis of facial image sequences using physical and anatomical models. IEEE Trans. Pattern Anal. Mach. Intell. **15**, 569–579 (1993)
14. Black, M., Yacoob, Y.: Tracking and recognizing rigid and non-rigid facial motions using local parametric models of image motion. In: International Conference on Computer Vision, pp. 374–381 (1995)

15. Blanz, V., Vetter, T.: Face recognition based on fitting a 3D morphable model. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 1063–1074 (2003)
16. Arulampalam, M., Maskell, A., Gordon, N., Clapp, T.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process. **50** (2002)
17. Dedeoglu, G., Baker, S., Kanade, T.: Resolution-aware fitting of active appearance models to low-resolution images. In: European Conference on Computer Vision (2006)
18. Yu, J., Bhanu, B., Xu, Y., Roy-Chowdhury, A.: Super-resolved facial texture under changing pose and illumination. In: International Conference on Image Processing (2007)

Face Variation

Carlos D. Castillo, David W. Jacobs
Department of Computer Science, University of Maryland, College Park, MD, USA

Synonym

Facial changes

Definition

Face variation refers to the way in which the appearance of the face changes due to changes in viewing conditions such as illumination or pose, or due to changes in properties of the face, such as its expression or age.

Introduction

Face recognition is a fundamental problem in biometrics. One of the chief sources of difficulty in face recognition is the large number of variations that can affect the appearance of faces. These include changes in lighting, pose, facial expression, makeup, hair, glasses, facial hair, occlusion by objects that block part of the face from view, aging, and weight gain or loss. Many studies suggest that these variations can significantly reduce the performance of recognition algorithms.

Some face recognition systems aimed at cooperative subjects deal with this problem by attempting to control these sources of variation. This may be appropriate for some applications. In these cases, pose can be controlled by requiring a subject to look into the camera, which is kept at a fixed height. Indoors, controlled lighting can be employed. And subjects may be requested to keep a neutral facial expression, and avoid variation in occluding objects, such as eye glasses or scarves. Working under such controlled conditions, face recognition systems have achieved high levels of accuracy [1].

However, in many cases large variations in appearance cannot be controlled. One may wish to recognize a person based on photographs taken some time before, as when one verifies that a person matches a passport photograph. In this case, changes in appearance due to aging, changes in weight, or variations in hair style will be inevitable. In many applications involving security or interactions between a computer or robot and a person, at least a few days may pass between the time a face is first learned and then later recognized. Even over short time periods there may be variation in a face due to changes in makeup, or in how recently a subject has shaved. Finally, in many applications, even lighting, pose or facial expression cannot be controlled, either because a subject is uncooperative or because one wishes to have the flexibility to recognize people as they move naturally through an environment, changing their position relative to the camera and lights.

There has been relatively less work on face recognition in the presence of these variations than for recognition under controlled conditions. Of these variations, lighting and pose variation have received the most attention. Many other sources of variation have been the subject of only a few research efforts. For example, to the authors' knowledge, there has been no work explicitly aimed at accounting for variations due to weight change. Furthermore, most research has been limited to the case in which conditions are controlled when subjects are enrolled into a *gallery* of known faces, so that, for example, all gallery images are taken in the same pose or lighting. Then, research focuses on matching a *probe* face viewed under different conditions to the correct entry in the gallery. Face recognition becomes much more difficult when the gallery is imaged under heterogenous conditions. For example, there may be a tendency to match a face viewed with side lighting to the gallery face of a different person viewed with the same lighting when the gallery face of the correct person is acquired under very different lighting conditions [2]. However, there are many applications, such as the organization of personal photos, in which it may not

be possible to acquire a gallery of images taken under controlled conditions.

Illumination

Changes in lighting can produce significant variation in the appearance of a face. These changes occur due to an actual variation in lighting, such as the difference between indoor and outdoor illumination, but they also occur when a person moves relative to even a fixed set of lights. Therefore, illumination changes must be accounted for in a wide range of recognition scenarios. Adini et al. [2] has shown that using common measures of image similarity, there is greater similarity between two images of different people taken under the same lighting conditions than between two images of the same person taken under quite different lighting conditions. As a consequence, in spite of a number of research efforts, existing recognition algorithms show much poorer performance in the presence of lighting variations than when used with controlled lighting [1].

A number of approaches have been taken in order to mitigate the effects of lighting change. Three common strategies include the following. First, one can apply image representations that are generically insensitive to lighting variation. Second, one can train a recognition system using sets of images that provide examples of the effects of lighting variation on images of faces. Third, one can use knowledge of the three-dimensional shape of faces either to predict the effect that changes in lighting might have on their appearance in images, or as a representation that is unaffected by lighting. These approaches are summarized briefly here; the reader can find more details in [3, 4, 5], and [6].

Determining the intrinsic properties of a scene independent of lighting conditions is a classic problem in computer vision that has been studied for decades. Multiplicative and additive effects of lighting can be removed by normalizing the mean and variance of the image intensities. ▶ Histogram equalization has been applied to remove lighting effects that produce a monotonic change in image intensities. Finally, some representations of images have been shown to be less sensitive to lighting variations, including the direction of image gradients, vectors containing the output of Gabor filters [7], or representations that attribute low frequency components of the image to lighting, and remove these effects. These and other, related techniques, have been shown to produce substantial improvements in recognition performance compared to methods that compare raw pixel intensities.

In a second approach, a *training* set of images is used to learn the effects of lighting variation on the appearance of faces. The training set may contain images of many individuals who are different from those the system will later try to recognize. These images show the variation in appearance of each person in the training set as the lighting varies. Methods such as ▶ Linear Discriminant Analysis may be used to then find representations of faces that best capture the information that varies between individuals, while discarding information that varies due to light, but not due to identity [8]. There is also a good deal of evidence that the set of images that a face produces under a wide range of lighting conditions occupy a low-dimensional linear subspace in the space of all possible images. This implies that when the gallery contains multiple images of each subject, taken under different lighting conditions, a linear subspace spanned by these images can be used to represent the subject.

A third set of methods makes use of knowledge of the 3D structure and surface reflectance properties of faces to predict and compensate for the effects of lighting [5]. This can involve obtaining a model of each face to be recognized. Acquisition systems that can capture the 3D structure of a face, along with the varying surface properties of eyebrows, lips, and skin exist. This makes the process of enrollment into the biometric system more complex, though. An alternative is to use general knowledge obtained from 3D scans of a training set of individuals other than the person to be recognized. In the latter case, a generic face model may be fit to a gallery image, producing a model specific to that person. A model of a person's face can be used to solve for the lighting that best matches that model to the probe image. Recognition can then be performed by comparing the probe image to a rendering of the model, produced by computer graphics. Other approaches may use the model to build representations of a face's appearance under diverse lighting conditions, and compare these to the probe image. Finally, if one obtains a 3D model as a probe, this can be directly compared to a 3D model acquired at enrollment.

Researchers have collected a number of data sets that contain images of the same individual under varying illumination, in order to measure the effect of lighting on recognition algorithms. Due to the difficulty of building such data sets, they are usually acquired with either a relatively small number of individuals or a small number of lighting conditions. For example, Carnegie Mellon University's Pose, Expression, and Illumination (PIE) data set contains images of 68 different people illuminated in turn by 21 different flash bulbs in known positions, while a variety of data sets contain images of more than a thousand individuals taken with just a few lighting conditions [9]. It is not clear how many lighting conditions are needed in a data set to thoroughly test recognition algorithms. The actual variability of lighting is very great, because even with lights distant from a face, the lighting intensity is a 2D function of direction. This means that it is difficult to record or simulate the lighting present in realistic conditions, and that it is also difficult to systematically explore the space of possible lighting conditions.

Pose

Face recognition with pose variation refers to recognizing faces when the cameras used to take gallery and probe images have different angles relative to the subject (e.g., Fig. 1). For example, there is a pose variation when subjects are described using a gallery of images taken with subjects facing the camera and when one uses a probe image of a subject seen in profile, but not when the probe image is simply taken from a different distance than the gallery images. When there is a pose variation, one may see different parts of the face in the gallery and probe images; for example, in a profile view, one side of the face may be unobserved. Moreover, the apparent size of different parts of the face may vary with pose. In profile, the cheek takes up a larger part of the image than it does when the face is viewed frontally, while the forehead may be more foreshortened. A number of experiments suggest that when one uses recognition algorithms that do not explicitly account for pose, performance deteriorates a great deal with significant pose variations.

Pose variations create a correspondence problem that does not occur with a number of other types of variations. It is common for general recognition algorithms to align faces by detecting and aligning a few features, such as the center of the eyes. When two images of a face are taken from frontal views, aligning the eyes tends to align all the other features of the face (although this is not quite true for some variations described below, such as changes in expression).

Face Variation. **Figure 1** The same person photographed in two different positions. Moving relative to the lighting and camera causes significant changes in appearance.

Many systems rely on this alignment by then comparing corresponding image pixels. However, when there is a pose variation, finding corresponding image pixels is much more difficult. Aligning the eyes in a frontal and a profile view will not align other parts of the face, such as the nose. Face recognition systems that can handle pose variation, then, must generally find some method of solving this correspondence problem.

This section discusses three approaches to this problem. The first involves representing multiple views of the face, so that a simple alignment with one of these views will match the probe. The second uses 2D image matching methods to find corresponding pixels. The third uses 3D representations to assist in solving for correspondence. These approaches are discussed further in [4, 5] and [6].

Many face recognition algorithms that are not designed to handle pose variation are still robust to small rotations of the head, of up to 15–30° [4]. This suggests that if the gallery contains images of each subject, taken at poses sampled by 30°, one of these gallery images will provide a good match to a probe. Such galleries have been constructed either by acquiring multiple images per subject, by constructing a 3D model of the subject and using it to generate appropriate views, using computer graphics, or by using training data to infer the changes of appearance in a face as viewpoint changes. These approaches may have the disadvantage of making enrollment into the gallery more complex, and may still degrade recognition performance to some degree when probes are taken at an angle between sampled directions.

Methods taking the second approach use some mechanism to find good correspondences between individual locations in the probe and gallery images. For example, [7] locates distinct features, such as the corner of the eyes, and builds descriptors of these locations using vectors containing the output of Gabor filters. Then corresponding features are matched between two images, allowing for changes in the relative position of features due to a pose change. Other work has matched individual pixels in images using ▶ optical flow [10] or stereo matching algorithms. These are matching methods developed for general computer vision problems in which a scene is viewed from different locations. These approaches may be supplemented by building a statistical model that captures the way a feature's appearance can vary with pose [4].

Finally, 3D face information may be used to account for pose. One way to do this is to acquire a 3D description of each subject when he or she is enrolled in the gallery. A small set of features can then be used to align this model with a 2D probe image. The 3D and 2D data must then be compared, which can be done, for example, by solving for the lighting that best matches them. Alternately, a system can obtain 3D information from the probe, and compare 3D representations directly. These approaches, though, depend on more complex sensing for enrollment, and possibly for recognition. An alternative approach (see [5]) builds a generic, 3D morphable model that can morph between the shapes of a set of training faces. This model can then be applied to any subject. By fitting the model to a probe image, the 3D structure of the probe face can be estimated. This can then be used to render the probe in a canonical pose, or it can be compared to similar 3D reconstructions of the gallery faces.

While progress has been made in handling pose variations, significant challenges remain. In particular, there are many applications in which one expects the gallery to contain a single image of the subject, and the probe to consist of a new image, taken in a new pose. For this problem, current methods have substantially worse performance than when pose is fixed between the probe and gallery. In addition, many methods for handling pose variation require substantially more computation than other methods, and can be very slow. This is partly because the process of finding a correspondence between the probe and gallery requires expensive optimization processes.

Expression and Occlusion

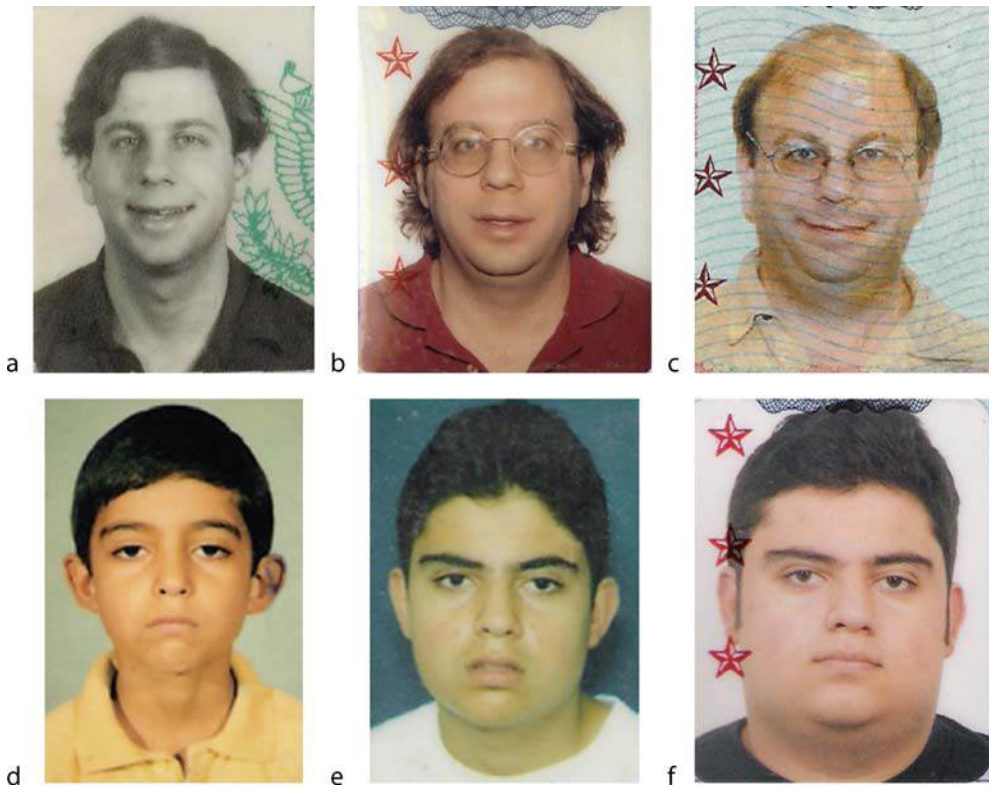
Changes in expression can also have a considerable effect on the appearance of a face that can have a major impact on recognition performance (see [11] for a fuller discussion). These can be divided into two sorts of effects. When one smiles, frowns, or purses one's lips, there is a change in shape, as the lips move and the cheeks alter their position. But expression can also cause facial features to appear or disappear. For example, smiling may reveal our teeth, blinking or winking may block an eye from view, frowning may cause new wrinkles to appear in the forehead. For this reason, it is convenient to class together changes in

expression with other occlusions, as when sunglasses or a scarf block part of the face from view.

Less work has been done on the problem of expression variation than on lighting or pose. One approach is to use recognition algorithms that can ignore or de-emphasize portions of the face that might be affected by expression change or occlusion. This can be done if training data is available that provides examples of these variations. Then, for example, Linear Discriminant Analysis can learn a linear projection that has the effect of placing less weight on portions of the face that are likely to change [8]. Or, one can divide the face into regions and learn weights that indicate the value of each region in identification. Regions of occlusion in a probe face can also be identified as regions that are not sufficiently similar to a space of face regions, and these can be discarded before matching the probe to the gallery [11]. In principal, changes in shape due to facial expression can be accounted for by using methods such as optical flow to find a correspondence between images of faces with different expressions [10]. However, such an approach must be able to distinguish between changes in shape caused by expression, and differences in shape between the faces of different people. Also, correspondences cannot be found when expression change causes features to appear or disappear. Because of their difficulty, many of the issues raised by changing expression have not been studied extensively.

Sources of Variation that Occur Over Time

Other sources of facial variation have received much less attention. These include changes in glasses, hair style, makeup, weight, or the effects of aging (See Fig. 2). While pose, expression and lighting can change from one moment to the next, these additional factors tend not to change very frequently. However, any system that wishes to recognize people after a period of a few months or a few years will have to account for these sources of variation.



Face Variation. **Figure 2** Two sets of photos showing changes in appearance over time. Left: passport photos taken at 10 year intervals. Right: photos taken at age 6, 16, and 23. There is a considerable change in appearance due to the effects of aging, weight gain, glasses, and changing hair and facial expression.

One reason that there may have been less effort directed at these variations is the difficulty of obtaining valid experimental data. For example, it would be daunting to collect images of large numbers of subjects before and after significant changes in weight. It is also much more challenging to collect face images over a period of many years than to collect images from different viewpoints, or with changes in lighting. The government does collect photos of individuals over long periods of time, for passports or drivers licenses, for example, but privacy concerns prevent widespread use of this data. As individuals post large collections of personal photos on the internet there is a growing opportunity to build innovative new data sets of face images, although by their nature, many of the imaging conditions in these photos are uncontrolled and unknown.

The complex set of factors that affect facial appearance over time are discussed in [12]. In children, there is significant change in face shape as they grow up. In adults, there is less change in shape due to aging, and more change in the appearance of skin due to exposure to sunlight and the appearance of wrinkles [12] and subsequent work describe experiments with a number of recognition algorithms, including two commercial systems, on data sets containing passport photos of nearly 2,000 individuals, with a time lag between photos ranging from 1–10 years. In a verification task that asks whether two photos come from the same or different people, performance is far below the levels achieved using photos taken under controlled conditions with little time lag [1]. It appears that there is a sharp increase in the difficulty of recognition when 1 year passes between images, and that, at least for adults, further passage of time, up to 10 years, creates only small additional increases in difficulty. It is not clear how much of these problems are due to aging, and how much can be attributed to other changes in, for example, weight or hair style that tend to occur over time, or even to other factors such as artifacts caused by the scanning of passport photos.

In addition to aging, a number of other sources of variation have been mentioned in the literature, but have not received much study. For example, a number of researchers have noted that the presence or absence of makeup on a face can affect the difficulty of recognizing it, but there is little systematic work in this area. Similarly, it is clear that significant changes in weight can affect facial appearance, but there has been little if any work in this area. Variations in hair style or grooming can also have a considerable affect on appearance; partially for this reason most approaches to face recognition focus on the inner part of the face, and attempt to ignore the outer head and hair. However, since the outer head and hair seem to be important in human face recognition, it seems that understanding hair appearance and its variations could be of potential value in face recognition systems.

Conclusions

In summary, while most work on face recognition has focused on settings in which there is little variation in a face or in the viewing conditions, there is also a growing amount of work that addresses face variations. In many cases, each source of variation has been addressed with methods specific to that type of variation. For example, lighting variation has been attacked using lighting insensitive image representations, while pose methods often focus on the correspondence problem. Two exceptions are first, model-based methods, such as those using morphable models, that extract a 3D model from an image, and then use computer graphics to normalize its appearance and remove face variations, and second, pattern recognition and learning methods, such as Linear Discriminant Analysis, that can potentially characterize any specific variation, provided there is appropriate training data.

Many face variations can cause significant degradation in performance in standard recognition methods. While interesting progress has been made in developing recognition methods that account for these variations, these methods generally still have performance that is substantially less than that can be achieved when variations are controlled.

While many challenges remain, this article has mentioned three in particular. First, there has been little research aimed at developing methods suitable for handling multiple simultaneous facial changes. For example, it is not clear whether many of the methods developed to handle lighting changes will be suitable when there is also pose variation. Second, most work has focused on situations in which the gallery images are taken under uniform conditions. Surely, for example, recognition will be more difficult when the gallery contains a single image of each person taken with different poses. Third, variations that occur over time

have not been well explored, and the relative importance of different effects, such as aging, weight change, or changing hair or makeup is not clear.

Acknowledgments

The authors have been supported by a fellowship from Apptis, Inc., and by a Honda Research Initiation Grant.

Related Entries

- ▶ Deformable Models
- ▶ Face Alignment
- ▶ Face Descriptors
- ▶ Face Pose Analysis
- ▶ Facial Expression Recognition
- ▶ Illumination Compensation

References

1. Phillips, P.J., Scruggs, W.T., O'Toole, A., Flynn, P., Bowyer, K., Schott, C., Sharpe, M.: FRVT 2006 and ICE 2006 large-scale results, National Institute of Standards and Technology Report NISTIR 7408 (2007)
2. Adini, Y., Moses, Y., Ullman, S.: Face recognition: The problem of compensating for changes in illumination direction. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 721–732 (1997)
3. Basri, R., Jacobs, D.: Illumination modeling for face recognition. In: Li, S., Jain, A. (eds.) The Handbook of Face Recognition, pp. 95–120. Springer, New York (2005)
4. Gross, R., Baker, S., Matthews, I., Kanade, T.: Face recognition across pose and illumination. In: Li, S., Jain, A. (eds.) The Handbook of Face Recognition, pp. 203–228. Springer, New York (2005)
5. Romdhani, S., Ho, J., Vetter, T., Kriegman, D.: Face recognition using 3-D models: Pose and illumination. Proc. IEEE **94**(11), 1977–1999 (2006)
6. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Comput. Surv. **35**(4), 399–458 (2003)
7. Lades, M., Vorbruggen, J., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R., Konen, W.: Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput. **42**(3), 300–311 (1993)
8. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)
9. Gross, R.: Face databases. In: Li, S., Jain, A. (eds.) The Handbook of Face Recognition, pp. 319–346. Springer, New York (2005)
10. Beymer, D., Poggio, T.: Image representations for visual learning. Science **272**, 1905–1909 (1996)
11. Martinez, A.: Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. IEEE Trans. Pattern Anal. Mach. Intell. **24**(6), 748–763 (2002)
12. Ramanathan, R., Chellappa, R.: Face verification across age progression. IEEE Trans. Image Process **15**(11), 3349–3361 (2006)

Face Verification

- ▶ Face Recognition, Overview
- ▶ Liveness Assurance in Face Authentication

Face Warping

- ▶ Image Warping
- ▶ Face Alignment
- ▶ Face Device
- ▶ Face Tracking

Face, Forensic Evidence of

Michael C. Bromby
Glasgow Caledonian University, Glasgow, UK

Synonyms

Face Identification; Face Recognition; Face Reconstruction; Facial Mapping

Definition

Using the face as a biometric feature requires an image or representation of the face, which is then subjected to manual or computerized analysis. This relies on an examination of the individual features (such as the eyes, nose, ears etc) or of the image as a whole (areas

of light and dark, texture, and color). Notably, face biometrics does not require active participation and can operate in a covert manner and may also be deployed from a distance. The major factors affecting the reliability and accuracy of face biometrics are illumination, pose, and expression.

Three main areas where the biometrics of the face is used are verification, identification, and reconstruction. First, the technique of comparing photographs of an offender with images of a suspect is occasionally termed ► *facial mapping*. Exculpatory evidence can be obtained if marked differences are apparent from expert analysis that cannot be explained. In contrast, similarity cannot indicate identity unless the presence of unique identifiers can be established. Second, computerized recognition and identification can provide a faster and more accurate method to search for a target in a database. This method may also be deployed in real time and generate a name or identifiable record when the target is present. Third, the process of skull reconstruction can provide a facial likeness, using either manual or computerized techniques, which is generally used for historical cases requiring identification.

Introduction

Face biometrics is regarded as less intrusive than other methods such as fingerprint analysis, iris scans, or palm morphologies, which generally require cooperation from the subject and an awareness of the procedure being undertaken. The face can be easily captured from a wide variety of low-cost sources, including public area CCTV, photographs, etc. The face also differs from the collectable forms of trail evidence left at the crime scene. Since the face is an intrinsic part of the owner, it cannot leave a physical trail other than a visual recording. Therefore, it may be seen as an exception to Locard's Exchange Principle in some instances.

Three main areas where the biometrics of the face is used are verification, identification, and reconstruction. Any of these methods may be susceptible to errors arising from the subjective nature of the interpretation of the face, whether by a person or by a machine. The ability of the face to move in three dimensions is also aggravated by internal movements of the eyes, lips, and cheek areas due to expressions such as smiling, frowning, blinking etc. Recognition from an image is also subject to the inherent limitations involved in general image processing such as resolution and lighting levels.

Evidence from facial biometrics is frequently seen as a corroborative tool to support other methods of identification. Its value as a single method of identification is lower than some other forms of biometric identifiers (notably fingerprints and DNA), and therefore, it is more often used in historical cases or those lacking in other forms of evidence.

Expert Image Analysis

Manual analysis of facial biometric features (often termed *facial mapping*) is primarily a feature of crime investigation. A variety of methods can be employed individually or in combination to compare a crime scene image of an offender with an image of the suspect. The face must be of a similar three-dimensional alignment as the head can move in a number of directions. Experts in this field have been providing reports to provide identification evidence, principally in criminal cases, since 1997 [1].

This category of facial biometrics employs a process of ► *one-to-one matching* – a verification process of checking allegations or suspicions, whereby other forms of evidence must be present to suggest the involvement of a specific individual. A one-to-one facial mapping technique provides additional scientific evidence to support the existing case. The use of this matching technique lends support to the judicial decision-makers who must be persuaded beyond reasonable doubt.

Facial Mapping Techniques

Although there are limited publications regarding facial image comparisons, many methodologies employed for image comparisons are drawn from peer-reviewed and accepted practices within other disciplines. Examples of video superimposition, morphological classifications, three-dimensional analysis, and geometric analysis are given in the literature and relate directly to forensic image comparisons for the purposes of identification [2–5].

The definition of "facial mapping" from the ACPO manual issued by the Working Group for Facial Identification is as follows:

1. Facial Identification by image comparison – concerned with the identity of an individual from scaled and aligned photographic images or by demonstrating morphologically comparable features, within a legal context.
2. To make a visual study of moving and/or still facial images in a variety of formats (video, digital, photographs etc) obtained from the scene of a crime or other source and make a scientific comparison with a suspect's facial image.
3. To present and demonstrate the significance of any area/point of similarity and difference, the presence or absence of a feature and any probability factor as well as the likelihood of repetition so as to formulate an opinion of similarity from these comparisons.
4. Similarities of features and facial proportion do not necessarily prove the identity, although differences may prove nonidentity. However, as the number of similarities increases, the number of people who share that particular combination of features/proportions decreases, thereby adding weight (to whatever degree) to the assumption that the persons in question are the same [6].

From this wide definition, it is clear that there is no single procedure or methodology required for comparing images. The ACPO document lists a number of methodologies that may be used to compare images. The document indicates that this list [6] is by no means exhaustive:

1. Drawn or electronically produced indicators/grids.
2. Transposed outlines (produced by hand or by computer).
3. Split or composite images (one or any portion of an image is overlaid on the second image to check/confirm correlation).
4. Video overlays/Wipes on a frame-by-frame basis.
5. Facial proportions/spatial distribution of features.

In general, these may be categorized into scaling and alignment methods to assess relative facial landmarks (size, shape, and position of facial features) and morphological comparisons. In practice, one or a combination of these methods is used, dependent on the imagery available. Additionally, ▶ *photogrammetry* has been employed using two images taken at different vantage points to create three-dimensional representations.

Appeal Court Cases

In reviewing the circumstances in which identification evidence based on CCTV or photographic imagery was admissible, the Court of Appeal for England and Wales identified four possible routes to achieve a valid identification. The fourth route was identified as being

> ▶ "a suitably qualified expert with facial mapping skills [who] can give opinion evidence of identification based on a comparison between images from the scene, (whether expertly enhanced or not) and a reasonably contemporary photograph of the defendant, provided the images and the photograph are available for the jury (*Stockwell* 97 Cr App R 260, *Clarke* [1995] 2 Cr App R 425 and *Hookway* [1999] Crim LR 750)" [7].

This response indicates that the admissibility of expert image analysis per se has remained unaltered since its first introduction as evidence of identification; and that the test of whether the court, in each instance, requires assistance in interpreting images through an expert witness is to be applied. In applying this test, the first route to achieving a valid identification as stated by the Vice President Rose, LJ declares that

> ▶ "where the photographic image is sufficiently clear, the jury can compare it with the defendant sitting in the dock (Dodson & Williams)" [7].

Under these circumstances, an expert opinion is clearly not required irrespective of whether the witness is indeed an expert. This scenario is significant as it can be distinguished from cases where an opinion is not admitted as evidence due to the lack of skill or knowledge claimed by the purported expert.

The *Attorney General's Reference* does, however, raise the question of what constitutes "suitable qualifications", which are not listed by the ACPO guidance, and presents investigators with a perennial problem. It could be seen that the Court of Appeal had the ideal opportunity to examine in greater detail the issues of admissibility, reliability, or indeed, sufficiency of image analysis as evidence of identification. Their reluctance to do so illustrates that there may not be a clear or singular answer to these issues.

In the UK case *R v Gray* criticisms were made by Mitting, J regarding the absence of statistical databases or any such means to determine a mathematical formula [8]. This did not develop any rule (as suggested in *R v Gardner*) that an expert cannot go further than saying "there are the following similarities",

leaving the ultimate decision to the jury, as opposed to the expert witness actually giving a view as to a degree of probability of the images being the same. The decision in *Gardner* does not doubt the admissibility of forensic image comparisons [9]. The appeal was based upon the inequality of arms as the defense team did not have access to the expert's laboratory material, upon which they could cross examine.

Mardia developed a database of facial statistics to determine whether, like fingerprints, there could be a certain number of matches on a face that would determine uniqueness. This study allows for a prevalence assessment of various facial feature classifications and angles of the face, although within a limited sample population of 358 Caucasian males [5]. Although this is not a nationally recognized database, it fulfills some criteria of objectivity within a measurement of uniqueness, as a sample size of only 50 achieved the same prevalence rates in a Home Office study by Wilcox [10].

Computer Analysis

As discussed earlier, an expert is able to make facial comparisons, using photographic evidence. The errors associated with human judgment may, on occasion, reduce the reliability of the expert and their evidence. Computerized facial recognition may eliminate the possible errors associated with both inter- and intraoperator variables. Many studies into computerized recognition have tried to adapt the psychological models of human recognition to work toward a fully computerized system of facial recognition.

► *Principal Component Analysis* is based on feature identification: a face is identified and stored, the image is then analyzed on the digital composition and the principal components or areas of light and dark are noted [11]. For example, thicker lips will possess a greater surface area and will vary in brightness and contrast between individuals. Areas of light and dark along the edge of the face also serve to identify face shape and relative size. A unique set of data for each individual face is created, which may then be used as a template or ► *eigenface* to enable the system to recognize the same face, or more correctly, the same set of data in the future.

An alternative model of ► *Graph Matching* [12] relies on the configurational identification of a face. This relates to the examination of the measurable distances between features and the relative ratios of height and width rather than the examination of the features themselves. The eyes can be identified automatically and the locations of the other features can be added if required by the software. A unique algorithm is created from the key points on the face; this algorithm is unique as a fingerprint or DNA profile. This second model is more similar to the task of facial mapping performed by experts, described earlier. However, with either method, there is still sufficient information to recognize and identify faces. The speed by which a result is obtained would favor an automatic computerized process, although it may be argued that a more thorough and reliable comparison can be made by using human input to locate the facial features.

Computerized techniques can assist ► *one-to-many identification* by searching through archive databanks of facial images. One-to-many matching for criminal justice purposes requires an extensive database of facial images collected either from police custody records or created from noncriminal records such as the face image held by the Passport Office or the Driver and Vehicle Licensing Authority (DVLA). A fully comprehensive national database of all adult facial images obtained from noncriminal records would not be in accordance with the protection offered under the legislation governing the use of data.

Reliability of Computerized Identification

The in-house testing of facial recognition systems by software companies can be extremely subjective, with varying aims and test data, depending on the actual use and requirements of the tasks that the algorithms were developed to perform. Accuracy and reliability can only be assessed by comparing a product with standardized references or samples and further analysis by independent bodies. The FERET Verification Testing Protocol for Face Recognition Algorithms was devised to provide an accurate and independent assessment of the reliability and accuracy of the existing facial recognition systems [13]. It also served to promote research in facial biometrics in academic and public/private sector industry, sponsored by the United States Department of Defense Counter-drug Technology Development Program. A Target set of "known individuals" and a Query set of "unknown faces" were presented to participating software developers. Two versions of testing were administered: the first assessed automatic facial location, and the second version

provided eye coordinates to assess the recognition performance of manual input systems. Enrollment and test data were collected according to strict guidelines to enable a fair comparison to be made. A scoring procedure was devised based on Receiver Operating Characteristic (ROC) graphs originally devised for SONAR false recognition rates [14].

A significant increase in performance was seen for the general field of facial biometric comparison and for each individual algorithm-based system [14]. Strengths and weaknesses of each algorithm were highlighted to facilitate further research to promote and improve the use of facial biometrics. It was evident from the FERET tests that further research was still required if facial biometrics were to compete with other forms of biometric identification such as fingerprints, even though progress had been made in these areas over a given period. A major fault of face recognition algorithms appeared to be sensitivity to variations in illumination, caused by the change in sunlight intensities throughout the day.

Automatic Recognition in Practice

By combining automatic recognition technology and criminal databases of known offenders, computer systems to alarm law enforcement agencies as to the real-time presence of a known criminal have been developed. The first CCTV and facial recognition system in the United Kingdom was instigated by the Metropolitan Police in Newham, East London [15]. In spite of the pressure from many civil liberties groups, the Mandrake system examined every passing face and alerted the police when an individual is recognized from the hit-list database. Despite analyzing every single face in a crowd, information was only stored when a match was made, and data from inconclusive analyses were discarded. The system relied wholly on a graph matching system, analyzing the area around the eyes and the nose, which was converted into an algorithm without any manual intervention. This means of crime prevention has inherent limitations, as unwarranted surveillance in anticipation of any crime occurring by chance is not permitted under Sections 28 and 29 of the Regulation of Investigatory Powers Act 2000. However, the selection of faces to be recognized and the specific locations of the CCTV cameras may permit facial recognition systems to be used for crime prevention.

By placing a surveillance system in a unique area and attaching a database specific to known criminals who would operate in that area, a reasonable successful hit rate can be achieved without infringing on the general privacy of the public. From the example put forward by Newham Council, other locations may be highlighted as target areas for particular types of offenders. Airports are prime examples of sites that are frequented by a variety of individuals involved in crimes ranging from terrorism to drug trafficking and illegal immigration. Security cameras are a regular feature of many public spaces and their presence has become ubiquitous because of their intrusive abilities to detect, recognize, and identify individuals without requiring an active participation or the knowledge of the subject.

The use of facial biometrics as a token for civilian verification of identity (for example, secure access, banking etc) is not so well employed. The benefits of not forgetting (as with passwords), not being lost (cards and keys), and being noninvasive (fingerprints etc) are often outweighed by higher false rejection rates when compared with other biometric systems.

Skull Reconstruction

In the absence of biological evidence such as DNA or identifiable personal artifacts, the naming of skeletalized or badly deformed remains may require the reconstruction of the face from the skull in order to identify the deceased.

Historically, the principle of relating the skeletal structure to the overlying soft tissue has been applied to all forms of reconstruction: 2-D drawings, 3-D clay sculpting, or computerized modeling. The skull clearly provides a vast amount of information on how the final face should appear. The sex, age, and racial origins can be determined, although any error will have significant repercussions throughout the whole procedure and will ultimately distort the reconstruction, possibly hindering the processes of recognition and identification by people familiar with the deceased. The relationship between hard and soft tissues of the face and facial tissue depth measurement provide the foundations for accurate reconstructions [16]. Some factors cannot be accounted for, such as the nutritional state of the individual.

While measuring the facial tissue depth, the number, and position of anthropometric landmarks are

subjective. Although published texts provide authoritative views, inter- and intraresearcher variation will persist in locating these points. Data from early cadaveric studies were subject to error due to shrinkage, bloating, and the effects of gravity when lying supine. Gravity, along with high radiation doses, persists to be a problem with modern advances in MRI and CT scans. Ultrasound is presented as the most reliable method, experiments providing comparative data from several ethnic groups [17, 18].

The Manchester Method relies on the knowledge of the gross anatomy of the face to recreate the muscle fibers and glands on a plaster cast copy of the skull [16] Unsurprisingly, this bottom-up process of rebuilding the face differs from the standard textbook descriptions of dissecting the facial musculature in a top-down fashion. Each muscle is created and attached, using published data and experience to recreate the underlying structures that will ultimately reflect the final skin surface with the minimum possibility of subjective interference.

In forensic cases, the addition of hairstyles, facial hair, blemishes, wrinkles, scars, or identifiable marks should not be added unless evidence suggests otherwise. Interestingly, details such as the hairline, forehead creases, eyelid patterns, nasolabial folds, and cheek shapes are some of the many features that can be determined, to some degree, from the skull and the previous muscle attachments. Creating a realistic and believable face is a difficult task balanced with the distraction of wrong information such as hair or eye color. Additional information may be superimposed using a computer software to generate a number of alternatives.

The accuracy of forensic facial reconstructions is the singularly most important factor in obtaining an identity for the deceased. Qualitative studies comparing the likeness with a photograph of the deceased have shown remarkable results. Blind testing using a variety of techniques has reported rates of 50, 65, and 75 per cent [16]. Quantitatively, very positive results have been obtained by conducting "identity parade" style face pools, using volunteers to assess the likeness against a number of targets [19]. The process of identifying unfamiliar faces is poorer than the ability to identify familiar faces, suggesting that these results are lower than what would be expected from family or friends of the reconstructed person.

Computerized face reconstruction, using three dimensional scanning of the skull has been reported as more reproducible than clay modeling, although subjectivity still remains in placing the pegs on the digitized skull [4, 20]. The benefit of a digital reconstruction is the flexibility of the final product, which may be aged or temporarily altered with greater ease than a more permanent clay final product.

Summary

The procedures involved in forensic face identification vary in both method and purpose according to whether the face is represented as a two-dimensional image or a three-dimensional skull. Evidence from all the three areas of expert or computer image analysis and skull reconstructions can be useful in obtaining an identification. The reliability and accuracy of each method may be prone to errors and the value of such evidence must be weighed in conjunction with other forms of identification or evaluated with some degree of caution if presented alone.

Related Entries

- ▶ Biometrics, Overview
- ▶ Biometric Recognition
- ▶ Face Recognition, Overview

References

1. *R v Ryan* (unreported): The Guardian Newspaper 26 April, 1991; cited in *R v Stockwell* (1997). Cr App R 260 at 264
2. Vanezis, P., Brierley, C.: Facial image comparison of crime suspects using video superimposition. Sci. Justice. **36**, 27–33 (1996)
3. Vanezis, M.P., Lu, D., Cockburn, J., Gonzalez, A., McCombe, G., Trujillo, O.: Morphological classification of facial features in adult Caucasian males based on an assessment of photographs of 50 subjects. J. Forensic Sci. **41**, 786–791 (1996)
4. Yoshino, M., Matsuda, H., Kubota, S., Imaizumi, K., Miyasaka, S.: Computer-assisted facial image identification system using a 3D physiognomic rangefinder. J. Forensic Sci. Int. **109**, 225–237 (2000)
5. Mardia, K., Coombes, A., Kirkbride, J., Linney A., Bowie, J.: On Statistical Problems with Face Identification from Photographs. J. Appl. Stat. **23**(6), 655–675 (1996)
6. Association of Chief Police Officers: National Working Practices in Facial Imaging. Home Office, London (2003) accessible at http://www.acpo.police.uk/asp/policies/Data/garvin_facial_imaging_guidelines.doc.
7. Rose, L.J.: In *Attorney-General's Reference No. 2 of 2002* (2003) Cr.App.R. 321

8. Mitting, J. in *R v Gray* (2003) EWCA 1001
9. *R v Gardner* (2004) EWCA 1639
10. Wilcox, R.: Facial Feature Prevalence Survey. London: Home Office Research, Development and Statistics Directorate. PRAS 33 (1994)
11. Pentland, A., Moshaddam, B., Starber, T.: View-based and modular Eigenfaces for face recognition. In: Proceedings of the IEEE Conference on Computerised Vision and Pattern Recognition 1994, pp. 84–91 (1993)
12. Wiskott, L., Fellous, J., Kruger, N., von der Malsburg, C.: Face recognition and gender determination. In: Proceedings of the International Workshop on Automatic Face and Gesture Recognition. Zurich (1995)
13. Phillips, P., Wechsler, H., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face recognition algorithms. Image Vis. Comput. J. **16**, 295–306 (1998)
14. Phillips, P., Moon, H., Rizvi, S., Rauss, P.: The FERET evaluation. In: Face Recognition from Theory to Applications. Springer, Berlin (1997)
15. Thomas, R.: As UK crime outstrips the US, a hidden eye is watching: Police switch on a camera that recognizes your face. The Observer, 11 October 1998, p. 5
16. Wilkinson, C.: Forensic Facial Reconstruction, Cambridge University Press, Cambridge (2004)
17. Auslebrooke, W.A., Becker, P.J., Iscan, M.Y.: Facial Soft Tissue Thickness in the Adult Male Zulu. Forensic Sci. Int. **79**, 83–102 (1996)
18. Lebedinskaya, G.U., Balueva, T.S., Veselovskaya, E.B.: Development of Methodological Principles for Reconstruction of the Face on the Basis of Skull Material, in Forensic Analysis of the Skull. pp. 183–198 Wiley-Liss, New York, (1993)
19. Wilkinson, C.M., Whittaker, D.K.: Skull Reassembly and the Implications for Forensic Facial Reconstruction. Sci. Justice **41**(3), 5–6 (2002)
20. Vanezis, P., Blowes, R.W., Linney, A.D., Tan, A.C., Richards, R., Neave, R.: Application of 3-D computer graphics for facial reconstruction and comparison with sculpting techniques. Forensic Sci. Int. **42**, 69–84 (1989)

Facial Action Coding

► Facial Expression Recognition

Facial Changes

► Face Variation

Facial Expression Analysis

► Facial Expression Recognition

Facial Expression Recognition

Maja Pantic
Department of Computing Imperial College London, London, UK

Synonyms

Facial Expression Analysis; Facial Action Coding

Definition

Facial expression recognition is a process performed by humans or computers, which consists of:

1. Locating faces in the scene (e.g., in an image; this step is also referred to as *face detection*),
2. Extracting facial features from the detected face region (e.g., detecting the shape of facial components or describing the texture of the skin in a facial area; this step is referred to as *facial feature extraction*),
3. Analyzing the motion of facial features and/or the changes in the appearance of facial features and classifying this information into some facial-expression-interpretative categories such as facial muscle activations like smile or frown, emotion (affect) categories like happiness or anger, attitude categories like (dis)liking or ambivalence, etc. (this step is also referred to as *facial expression interpretation*).

Introduction

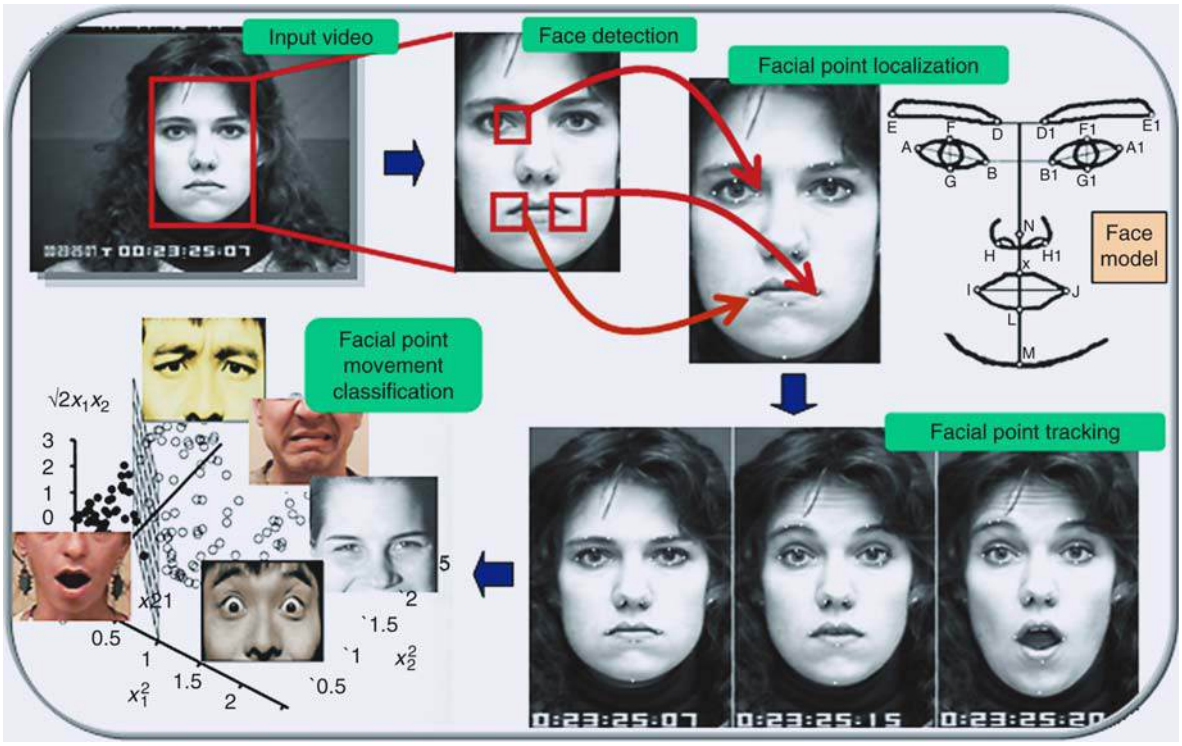
A widely accepted prediction is that computing will move to the background, weaving itself into the fabric of our everyday living and projecting the human user into the foreground. To realize this goal, next-generation computing (a.k.a. pervasive computing,

ambient intelligence, and ▶ human computing) will need to develop human-centered ▶ user interfaces that respond readily to naturally occurring, multimodal, human communication [1]. These interfaces will need the capacity to perceive and understand intentions and emotions as communicated by social and affective signals. Motivated by this vision of the future, automated analysis of nonverbal behavior, and especially of facial behavior, has attracted increasing attention in computer vision, pattern recognition, and human-computer interaction [2–5]. To wit, facial expression is one of the most cogent, naturally preeminent means for human beings to communicate emotions, to clarify and stress what is said, to signal comprehension, disagreement, and intentions, in brief, to regulate interactions with the environment and other persons in the vicinity [6, 7]. Automatic analysis of facial expressions forms, therefore, the essence of numerous next-generation-computing tools including ▶ affective computing technologies (proactive and affective user interfaces), learner-adaptive tutoring systems, patient-profiled personal wellness technologies, etc.

The Process of Automatic Facial Expression Recognition

The problem of machine recognition of human facial expression includes three subproblem areas (Fig. 1): (1) finding faces in the scene, (2) extracting facial features from the detected face region, (3) analyzing the motion of facial features and/or the changes in the appearance of facial features, and classifying this information into some facial-expression-interpretative categories (e.g., emotions, facial muscle actions, etc.).

The problem of *finding faces* can be viewed as a segmentation problem (in machine vision) or as a detection problem (in pattern recognition). It refers to identification of all regions in the scene that contain a human face. The problem of finding faces (*face localization*, *face detection*) should be solved regardless of clutter, occlusions, and variations in head pose and lighting conditions. The presence of non-rigid movements due to facial expression and a high degree of variability in facial size, color and texture make this problem even more difficult. Numerous techniques have been developed for face detection in still images [8, 9],



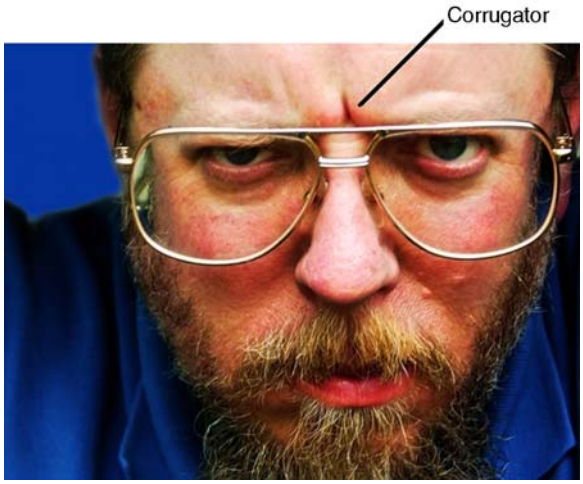
Facial Expression Recognition. Figure 1 Outline of an automated, geometric-features-based system for facial expression recognition (for details of this system, see [4]).

(see ► Face Localization). However, most of them can detect only upright faces in frontal or near-frontal view. Arguably the most commonly employed face detector in automatic facial expression analysis is the real-time face detector proposed by Viola and Jones [10].

The problem of feature extraction can be viewed as a dimensionality reduction problem (in machine vision and pattern recognition). It refers to transforming the input data into a reduced representation set of features which encode the relevant information from the input data. The problem of *facial feature extraction* from input images may be divided into at least three dimensions [2, 4]: (1) Are the features holistic (spanning the whole face) or analytic (spanning subparts of the face)?; (2) Is temporal information used?; (3) Are the features view- or volume based (2-D/3-D)?. Given this glossary, most of the proposed approaches to facial expression recognition are directed toward static, analytic, 2-D facial feature extraction [3, 4]. The usually extracted facial features are either *geometric features* such as the shapes of the facial components (eyes, mouth, etc.) and the locations of facial fiducial points (corners of the eyes, mouth, etc.), or *appearance features* representing the texture of the facial skin in specific facial areas including wrinkles, bulges, and furrows. Appearance-based features include learned image filters from Independent Component Analysis (ICA), Principal Component Analysis (PCA), Local Feature Analysis (LFA), Gabor filters, integral image filters (also known as box-filters and Haar-like filters), features based on edge-oriented histograms, etc, (see ► Skin Texture, and ► Feature Extraction). Several efforts have also been reported which use both geometric and appearance features (e.g., [3]). These approaches to automatic facial expression analysis are referred to as *hybrid* methods. Although it has been reported that methods based on geometric features are often outperformed by those based on appearance features using, e.g., Gabor wavelets or eigenfaces, recent studies show that in some cases geometric features can outperform the appearance-based ones [4, 11]. Yet, it seems that using both geometric and appearance features might be the best choice in the case of certain facial expressions [11].

Contractions of facial muscles, which produce facial expressions, induce movements of the facial skin and changes in the location and/or appearance of facial features (e.g., contraction of the Corrugator muscle induces a frown and causes the eyebrows to move towards each other, usually producing wrinkles between the eyebrows; Fig. 2). Such *changes can be detected* by analyzing optical flow, facial-point- or facial-component-contour-tracking results, or by using an ensemble of classifiers trained to make decisions about the presence of certain changes (e.g., whether the nasolabial furrow is deepened or not) based on the passed appearance features. The optical flow approach to describing face motion has the advantage of not requiring a facial feature extraction stage of processing. Dense flow information is available throughout the entire facial area, regardless of the existence of facial components, even in the areas of smooth texture such as the cheeks and the forehead. Because optical flow is the visible result of movement and is expressed in terms of velocity, it can be used to represent directly the facial expressions. Many researchers adopted this approach [2, 3]. Until recently, standard optical flow techniques were, arguably, most commonly used for tracking facial characteristic points and contours as well [4]. In order to address the limitations inherent in optical flow techniques such as the accumulation of error and the sensitivity to noise, occlusion, clutter, and changes in illumination, recent efforts in automatic facial expression recognition use sequential state estimation techniques (such as Kalman filter and Particle filter) to track facial feature points in image sequences (e.g., [4, 11]).

Eventually, dense flow information, tracked movements of facial characteristic points, tracked changes in contours of facial components, and/or extracted



Facial Expression Recognition. Figure 2 Facial appearance of the Corrugator muscle contraction (coded as in the FACS system, [14]).

appearance features are translated into a description of the displayed facial expression. This description (*facial expression interpretation*) is usually given either in terms of shown affective states (emotions) or in terms of activated facial muscles underlying the displayed facial expression. This stems directly from two major approaches to facial expression measurement in psychological research [12]: message and sign judgment. The aim of message judgment is to infer what underlies a displayed facial expression, such as affect or personality, while the aim of sign judgment is to describe the "surface" of the shown behavior, such as facial movement or facial component shape. Thus, a brow frown can be judged as "anger" in a message-judgment and as a facial movement that lowers and pulls the eyebrows closer together in a sign-judgment approach. While message judgment is all about interpretation, sign judgment attempts to be objective, leaving inference about the conveyed message to higher order decision making. Most commonly used facial expression descriptors in message judgment approaches are the six basic emotions (fear, sadness, happiness, anger, disgust, surprise; see Fig. 3) proposed by Ekman and discrete emotion theorists [13], who suggest that these emotions are universally displayed and recognized from facial expressions. Most commonly used facial action descriptors in sign judgment approaches are the Action Units (AUs) defined in the Facial Action Coding System (FACS; [14]). Most facial expressions analyzers developed, so far, target human facial affect analysis and attempt to recognize a small set of prototypic emotional facial expressions like happiness and anger [2, 5]. However, several promising prototype systems were reported that can recognize deliberately produced AUs in face images and even few attempts towards recognition of spontaneously displayed AUs have been recently reported as well [3–5]. While the older methods employ simple approaches including expert rules and machine learning methods such as neural networks to classify the relevant information

Facial Expression Recognition. **Figure 3** Prototypic facial expressions of six basic emotions (left-to-right from top row): disgust, happiness, sadness, anger, fear, and surprise.

from the input data into some facial-expression-interpretative categories, the more recent (and often more advanced) methods employ probabilistic, statistical, and ensemble learning techniques, which seem to be particularly suitable for automatic facial expression recognition from face image sequences [3, 5].

Evaluating Performance of an Automated System for Facial Expression Recognition

The two crucial aspects of evaluating performance of a designed automatic facial expression recognizer are the utilized training/test dataset and the adopted evaluation strategy.

Having enough labeled data of the target human facial behavior is a prerequisite in designing robust automatic facial expression recognizers. Explorations of this issue showed that, given accurate 3-D alignment of the face (see ► Face Alignment), at least 50 training examples are needed for moderate performance (in the 80% accuracy range) of a machine-learning approach to recognition of a specific facial expression [4]. Recordings of spontaneous facial behavior are difficult to collect because they are difficult to elicit, short lived, and filled with subtle context-based changes. In addition, manual labeling of spontaneous facial behavior for ground truth is very time consuming, error prone, and expensive. Due to these difficulties, most of the existing studies on automatic facial expression recognition are based on the "artificial" material of deliberately displayed facial behavior, elicited by asking the subjects to perform a series of facial expressions in front of a camera. Most commonly used, publicly available, annotated datasets of posed facial expressions include the Cohn-Kanade facial expression database, JAFFE database, and MMI facial expression database [4, 15]. Yet, increasing evidence suggests that deliberate (posed) behavior differs in appearance and timing from that which occurs in daily life. For example, posed smiles have larger amplitude, more brief duration, and faster onset and offset velocity than many types of naturally occurring smiles. It is not surprising, therefore, that approaches that have been trained on deliberate and often exaggerated behaviors usually fail to generalize to the complexity of expressive behavior found in real-world settings. To address the general lack of a reference set of (audio and/or) visual recordings of human spontaneous behavior, several efforts aimed at development of such datasets have been recently reported. Most commonly used, publicly available, annotated datasets of spontaneous human behavior recordings include SAL dataset, UT Dallas database, and MMI-Part2 database [4, 5].

In pattern recognition and machine learning, a common evaluation strategy is to consider correct classification rate (*classification accuracy*) or its complement error rate. However, this assumes that the natural distribution (prior probabilities) of each class are known and balanced. In an imbalanced setting, where the prior probability of the positive class is significantly less than the negative class (the ratio of these being defined as the *skew*), accuracy is inadequate as a performance measure since it becomes biased towards the majority class. That is, as the skew increases, accuracy tends towards majority class performance, effectively ignoring the recognition capability with respect to the minority class. This is a very common (if not the default) situation in facial expression recognition setting, where the prior probability of each target class (a certain facial expression) is significantly less than the negative class (all other facial expressions). Thus, when evaluating performance of an automatic facial expression recognizer, other performance measures such as *precision* (this indicates the probability of correctly detecting a positive test sample and it is independent of class priors), *recall* (this indicates the fraction of the positives detected that are actually correct and, as it combines results from both positive and negative samples, it is class prior dependent), *F1-measure* (this is calculated as *2*recall*precision/(recall + precision)*), and ROC (this is calculated as $P(x|positive)/P(x|negative)$, where $P(x|C)$ denotes the conditional probability that a data entry has the class label C, and where a ROC curve plots the classification results from the most positive to the most negative classification) are more appropriate. However, as a confusion matrix shows all of the information about a classifier's performance, it should be used whenever possible for presenting the performance of the evaluated facial expression recognizer.

Applications

The potential benefits from efforts to automate the analysis of facial expressions are varied and numerous and span fields as diverse as cognitive sciences, medicine, communication, education, and security [16].

When it comes to computer science and computing technologies, facial expressions provide a way to communicate basic information about needs and demands to the machine. Where the user is looking (i.e., gaze tracking) can be effectively used to free computer users from the classic keyboard and mouse. Also, certain facial signals (e.g., a wink) can be associated with certain commands (e.g., a mouse click) offering an alternative to traditional keyboard and mouse commands. The human capability to "hear" in noisy environments by means of lip reading is the basis for bimodal (audiovisual) speech processing (see Lip-Movement Recognition), which can lead to the realization of robust speech-driven *user interfaces*. To make a believable *talking head* (avatar) representing a real person, recognizing the person's facial signals and making the avatar respond to those using synthesized speech and facial expressions is important. Combining facial expression spotting with facial expression interpretation in terms of labels like "did not understand", "disagree", "inattentive", and "approves" could be employed as a tool for monitoring human reactions during videoconferences, web-based lectures, and automated tutoring sessions. The focus of the relatively, recently initiated research area of *affective computing* lies on sensing, detecting and interpreting human affective states (such as pleased, irritated, confused, etc.) and devising appropriate means for handling this affective information in order to enhance current ▶ HCI designs. The tacit assumption is that in many situations human-machine interaction could be improved by the introduction of machines that can adapt to their users and how they feel. As facial expressions are our direct, naturally preeminent means of communicating emotions, machine analysis of facial expressions forms an indispensable part of affective HCI designs.

Monitoring and interpreting facial expressions can also provide important information to lawyers, police, security, and intelligence agents regarding *person's identity* (research in psychology suggests that facial expression recognition is much easier in familiar persons because it seems that people display the same, "typical" patterns of facial behaviour in the same situations), *deception* (relevant studies in psychology suggest that visual features of facial expression function as cues to deception), and *attitude* (research in psychology indicates that social signals including accord and mirroring – mimicry of facial expressions, postures, etc., of one's interaction partner – are typical, usually unconscious gestures of wanting to get along with and be liked by the interaction partner). Automated facial reaction monitoring could form a valuable tool in law enforcement, as now only informal interpretations are typically used. Systems that can recognize friendly faces or, more importantly, recognize unfriendly or aggressive faces and inform the appropriate authorities represent another application of facial measurement technology.

Concluding Remark

Faces are tangible projector panels of the mechanisms which govern our emotional and social behaviors. The automation of the entire process of facial expression recognition is, therefore, a highly intriguing problem, the solution to which would be enormously beneficial for fields as diverse as medicine, law, communication, education, and computing. Although the research in the field has seen a lot of progress in the past few years, several issues remain unresolved. Arguably the most important unattended aspect of the problem is how the grammar of facial behavior can be learned (in a human-centered, context-profiled manner) and how this information can be properly represented and used to handle ambiguities in the observation data. This aspect of machine analysis of facial expressions forms the main focus of the current and future research in the field.

Related Entries

- ▶ Face Alignment
- ▶ Face Localization
- ▶ Feature Extraction
- ▶ Lip Movement Recognition
- ▶ Skin Texture

References

1. Pantic, M., Pentland, A., Nijholt, A., Huang, T.S.: Human computing and machine understanding of human behavior: A Survey. Lect. Notes Artif. Intell. **4451**, 47–71 (2007)
2. Pantic, M., Rothkrantz, L.J.M.: Toward an affect-sensitive multimodal HCI. Proceedings of the IEEE **91**(9), 1370–1390 (2003)
3. Tian, Y.L., Kanade, T., Cohn, J.F.: Facial expression analysis. In: Li, S.Z., Jain, A.K. (eds.) Handbook of Face Recognition, pp. 247–276. Springer, New York (2005)

4. Pantic, M., Bartlett, M.S.: Machine analysis of facial expressions. In: Delac, K., Grgic, M. (eds.) 1Face Recognition, pp. 377–416. I-Tech Education and Publishing, Vienna, Austria (2007)
5. Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S.: A survey of affect recognition methods: Audio, visual, and spontaneous expressions. IEEE Trans. Pattern Anal. Mach. Intell. **31**(1), 39–58 (2009)
6. Ambady, N., Rosenthal, R.: Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. Psychol. Bull. **111**(2), 256–274 (1992)
7. Ekman, P., Rosenberg, E.L. (eds.): What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system. Oxford University Press, Oxford, UK (2005)
8. Yang, M.H., Kriegman, D.J., Ahuja, N.: Detecting faces in images: A survey. IEEE Trans. Pattern Anal. Mach. Intell. **24**(1), 34–58 (2002)
9. Li, S.Z., Jain, A.K. (eds.): Handbook of face recognition. Springer, New York (2005)
10. Viola, P., Jones, M.: Robust real-time face detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
11. Pantic, M., Patras, I.: Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences. IEEE Trans. Syst. Man Cybern. B Cybern. **36**(2), 433–449 (2006)
12. Cohn, J.F., Ekman, P.: Measuring facial actions. In: Harrigan, J.A., Rosenthal, R., Scherer, K. (eds.) The New Handbook of Methods in Nonverbal Behavior Research, pp. 9–64. Oxford University Press, New York (2005)
13. Keltner, D., Ekman, P.: Facial expression of emotion. In: Lewis, M., Haviland-Jones, J.M. (eds.) Handbook of Emotions, pp. 236–249. Guilford Press, New York (2000)
14. Ekman, P., Friesen, W.V., Hager, J.C.: Facial action coding system. A Human Face, Salt Lake City, USA (2002)
15. Pantic, M., Valstar, M.F, Rademaker, R., Maat, L. Web-based database for facial expression analysis. Proc. IEEE Int'l Conf. Multimedia & Expo (ICME) 317–321 (2005)
16. Ekman, P., Huang, T.S., Sejnowski, T.J., Hager, J.C. (eds.): NSF Understanding the Face. A Human Face eStore, Salt Lake City, USA, (see Library) (1992)

Facial Landmarks

A number of pixels in a face image clearly corresponds to some extract physiological semantics, such as the eye corners, eye centers, mouth corners, nose tips, etc. These feature points are called facial landmarks. They are generally used to align different face images for accurate matching.

▶ Face Misalignment Problem

Facial Mapping

Facial mapping is a frequently used term to describe one-to-one matching of crime scene and suspect images undertaken by an expert. A number of different methods may be used in combination to compare two images to support the comparison. Also, a number of comparisons may be made of the face from different angles if multiple images are available from each source.

▶ Face, Forensic Evidence of

Facial Motion Estimation

▶ Face Tracking

Facial Photograph

▶ Photography for Face Image Data

Factor Analysis

▶ Session Effects on Speaker Modeling

Failure to Acquire Rate

Both the acquiring conditions and the flaw of biometric itself may cause failure to acquire a biometric trait. The percentage of this failure is defined as "Failure to Acquire Rate." For instance, very low quality face image may cause the failure of face detection and subsequent feature extraction.

▶ Evaluation of Biometric Quality Measures
▶ Performance Evaluation, Overview

Failure-to-Enrol Rate

Failure-to-enrol rate is defined as the proportion of enrollment transactions in which zero instances were enrolled. Enrollment in one or more instances is considered to be successful in the case, the systems accept multiple biometric samples per person.

► Finger Vein Reader

Fake Finger Detection

► Anti-spoofing
► Fingerprint Fake Detection

False Match Rate

The probability that a biometric system will indicate that two biometric templates match although they are not derived from the same individual and should not match.

► Fingerprint Image Quality
► Iris on the Move

False Negative Rate

False Negative Rate means that how many percentages of the authentic test samples are incorrectly classified as the imposter class. Take the example of the computer account login system, False Negative Rate means how many percentages of legal users are recognized as illegal users. As one can see immediately, False Positive Rate and False Negative Rate are two metrics that counter each other. For any given biometrics modality with given matching algorithm, requirement of low False Positive Rate would unavoidably bring high False Negative Rate, and vice versa. Performance comparison between different algorithms is usually done by comparing False Negative Rate at a fixed False Positive Rate.

► Biometric System Design, Overview
► Iris Recognition, Overview

False Non-Match Rate

False non-match rate is the proportion of genuine comparisons that result in false non-match. False non-match is the decision of non-match when comparing biometric samples that are from same biometric source (i.e., genuine comparison).

► Biometric System Design, Overview
► Fingerprint Image Quality
► Iris on the Move

False Positive Rate

False Positive Rate means how many percentage of the imposter test samples are incorrectly classified as the authentic class. For example, in a computer account login system, False Positive Rate is what percentage of the illegal users recognized as legal users. In applications, which require high security, False Positive Rate is always required to be as small as possible.

► Biometric System Design, Overview
► Iris Recognition, Overview

Feathering

Feathering is a feature which occurs on the outsole as a result of an abrasive wear and has some resemblance to the ridge characteristics and bifurcations of fingerprint patterns. It is the result of frictional abrasive forces applied to the outsole surface such as when scuffing

or dragging the shoe. This feature is also known as a Schallamach pattern.

► Footwear Recognition

Feature Detection

Finding significant features in images such as landmarks, edges, or curves. For example, a facial feature detector aims to find the positions of the center of an eye, the corners of a mouth, or the top of a nose in a face image. In the case of an iris image, features may mean the edges inside an iris or the boundaries around the iris.

► Iris Segmentation Using Active Contours

Feature Extraction

► Biometric Algorithms

Feature Fusion

Producing a merged feature vector from a set of feature vectors representing different aspects of biometric data. The data can originate from different sensors, and also from different properties of a signal that originate from the same sensor.

► Fusion, Feature-Level
► Multiple Experts

Feature Map

The image produced from a target image to enhance the signals of a particular type, such as edges, ridges, or valleys is referred as a "feature map." Face alignment programs typically rely on statistics computed from such features to distinguish facial features from other regions of the image. More sophisticated feature maps can be constructed to capture complicated local image structures and enhance the stability.

► Face Alignment

Feature Selection

Feature selection techniques are aimed towards finding an optimal feature set for a specific purpose, such as the optimization of a biometric system verification performance. In general, feature selection algorithms try to avoid the evaluation of all the possible feature combinations when searching for an optimal feature vector, since these grow exponentially as the number of feature increases.

► Signature Features

Feature Vector

Feature vector is a multidimensional vector that is obtained from a face by using feature extraction and image processing techniques to be used and that is used to memorize and recognize the face.

► Face Databases and Evaluation

Features

Biometric features are the information extracted from biometric samples which can be used for comparison with a biometric reference. For example, characteristic measures extracted from a face photograph such as eye distance or nose size etc. The aim of the extraction of biometric features from a biometric sample is to

remove superfluous information which does not contribute to biometric recognition. This enables a fast comparison and an improved biometric performance, and may have privacy advantages.

► Biometric Algorithms
► Vascular Image Data Format, Standardization

Features vs. Templates

► Face Recognition, Geometric vs. Appearance-Based

Fidelity

The degree of similarity between a biometric sample and its source. Fidelity of a sample is comprised of individual components of fidelity attributed to each step through which it is processed (e.g., compression).

► Biometric Sample Quality
► National Institute for Standards and Technology

Field of View (FOV)

Field of view (FOV) is the angular portion of visible space which is comprised into the image region. The FOV of the human eye is around 150°. The camera FOV depends both on the size of the camera sensor and the geometry of the lens. The camera focal length determines the field of view falling within the sensor area, thus determining also the magnification factor of the image. A shorter camera focal length produces a wider FOV, while a longer focal length produces a smaller FOV.

► Face Device

Finger Data Interchange Format, Standardization

Raul Sanchez-Reillo[1], Robert Mueller[2]
[1]University Carlos III of Madrid, Avda. Universidad, Leganes (Madrid), Spain
[2]Giesecke & Devrient GmbH, Prinzregentenstr. Muenchen, Germany

Synonyms

Encoded finger data; Fingerprint data interchange format

Definition

Set of ISO Standards that define common formats to encode information related to finger-based biometrics. Those formats are defined to allow interoperability among different vendors worldwide, and have been developed by the international community taking part in ISO/IEC JTC1/SC37 standardization subcommittee. Those documents define not only the way a fingerprint image has to be encoded, but also the way a feature vector composed of ► minutiae points has to be stored and/or transmitted. Furthermore, formats for the ► spectral data of the finger, as well as its skeletal data are defined.

Introduction

Standardization is essential for the wide-spread adoption of technologies in open mass applications. Fingerprint recognition is not only the most prominent biometric measure, but also the biometric trait with the largest databases and the best long-term experience. Fingerprints are used in applications such as physical access control and digital signature creation but also national ID card schemes and other governmental projects. The need for standardization is conspicuous in every single area where it is not applied.

The SC37 Subcommittee from ISO/IEC JTC1 deals with the standardization of biometrics. Among the many aspects of its work, SC37's Working Group 3 is devoted to defining Interchange Data Formats for a variety of biometric modalities. To accomplish this,

a multipart standard is under development, covering several biometric modalities. Such multipart standard is known as ISO/IEC 19794. There are four parts in this standard which cover finger-based biometrics, or what can be better understood as fingerprint biometrics.

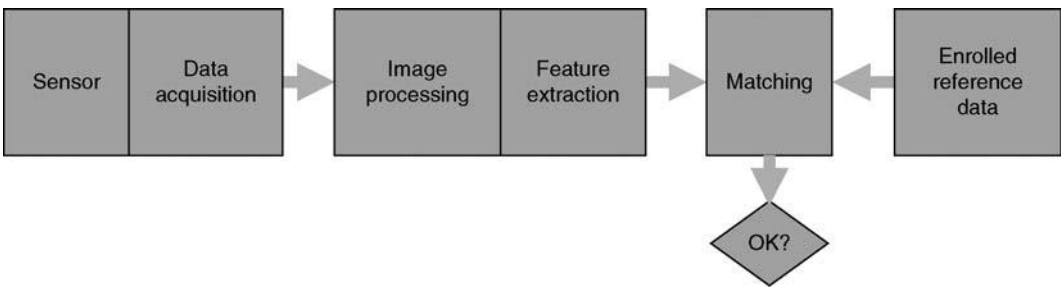
1. Part 2 of the Standard series, deals with the way a minutiae-based feature vector or template has to be coded
2. Part 3 standardizes the way to code information referring to the spectral information of the fingerprint
3. Part 4 determines the coding of a fingerprint raw image and
4. Part 8 establishes a way to code a fingerprint by its skeleton

Figure 1 shows the basic architecture of a typical fingerprint verification system. A finger is presented to a sensor and a raw image acquired. Image processing techniques enhance the image quality before a feature vector of characteristic features can be extracted. The features are compared with a previously recorded reference data set to determine the similarity between the two sets before the user presenting the finger is authenticated. The reference data is stored in a database or on a portable data carrier.

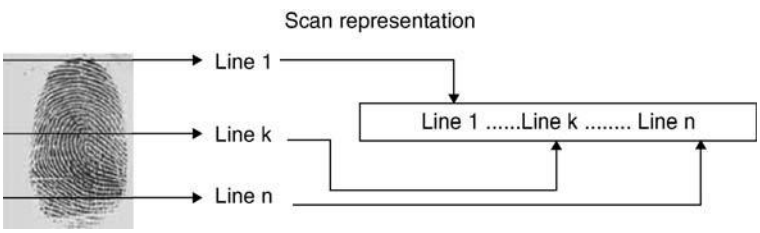
The following subsections explain the basic characteristics of each type of finger-based standard. The image standard (Part 4) is presented first as it is the first step in the fingerprint comparison process as shown in the architecture above. This is followed by the other finger based standards, each of which deals with samples already processed.

Finger Images

As already mentioned, the way a fingerprint image is to be coded is defined in ISO/IEC 19794-4 International Standard [1], whose title is "Information technology - Biometric data interchange formats - Part 4: Finger image data." The way the finger is scanned is out of the scope of the standard, but after image acquisition, the image shall represent a finger in upright position, i.e., vertical and with the tip of the finger in the upper part of the image. The way to code such an image is represented in Fig. 2, where the top line is the first to be stored and/or transmitted. This is in contradiction to mathematical graphing practice but in conjunction with typical digital image processing. For those images that require two or more bytes per pixel intensity, the most significant byte is stored/transmitted first, and bytes follow most significant bit coding.



Finger Data Interchange Format, Standardization. **Figure 1** Typical Biometric Verification System.



Finger Data Interchange Format, Standardization. **Figure 2** Coding structure of a fingerprint image. Image taken from [1].

This International Standard also includes a set of constraints for image acquisition. It determines the pixel aspect ratio, which shall be between 0.99 and 1.01 (horizontal/vertical sizes), as well as several image acquisition levels, as stated in Table 1.

After the requirements for the image to be stored or transmitted have been specified, this International Standard details the structure of the data record referring to a finger image. Following CBEFF specifications [2] (see entry "Common Biometric Exchange Framework Formats"), a record referring to a finger image has the following structure (for details refer to the last version of this International Standard [1]):

- A single fixed-length (32-byte) general record header containing information about the overall record, with the following fields:
  - Format identifier (4 bytes with the hexadecimal value 0x46495200) and version number (coded in another 4 bytes)
  - Record length (in bytes) including all finger images within that record (coded in 6 bytes)
  - Capture device ID (2 bytes) and Image acquisition level (2 bytes)
  - Number of fingers (1 byte), Scale units used (1 byte), and Scan resolution used (2 bytes for horizontal and another 2 for vertical resolution)
  - Image resolution, coded the same way as the scan resolution, and whose value shall be less or equal to scan resolution
  - Pixel depth (1 byte) and Image compression algorithm used (coded in 1 byte)
  - 2 bytes reserved for future use
- A single finger record for each finger, view, multi-finger image, or palm consisting of:
  - A fixed-length (14-byte) finger header containing information pertaining tothe data for a single or multi-finger image, which gives information about:
    - Length of the finger data block (4 bytes)
    - Finger/palm position (1 byte)
    - Count of views (1 byte) and View number (1 byte)
    - Finger/palm image quality (1 byte) and Impression type (1 byte)
    - Number of pixels per horizontal line (2 bytes) and Number of horizontal lines (2 bytes)
    - 1 byte reserved for future use
  - Compressed or uncompressed image data view for a single, multi-finger, or palm image, which has to be smaller than $43x10^8$ bytes.

Finger Data Interchange Format, Standardization. Table 1 Image acquisition levels for finger biometrics. Extract from Table 1 in [1]

Setting level	Scan resolution (dpi)	Pixel depth (bits)	Gray levels
10	125	1	2
20	250	3	5
30	500	8	80
31	500	8	200
35	750	8	100
40	1,000	8	120
41	1,000	8	200

The raw finger format is used, for example, in databases containing standard fingerprints. Law enforcement agencies are typical applicants of the standard. The largest fingerprint image databases are maintained by the FBI in the United States and are encoded with a national counterpart of this standard.

Fingerprint Minutiae

While Part 4 of the 19794 Series of Standards is dedicated to raw biometric sample data, Part 2 refers to the format in which a minutiae-based feature vector or template has to be coded. Therefore ISO/IEC 19794-2 "Information Technology - Biometric data interchange Formats - Part 2: Finger minutiae data" [3] deals with processed biometric data, ready to be sent to a comparison block to obtain a matching score.

Finger minutiae are local point patterns present in a fingerprint image. The comparison of these characteristic features is sufficient to positively identify a person. Sir Francis Galton first defined the features of a fingerprint [4].

In order to reach interoperability, this International Standard defines not only the record format, but also the rules for fingerprint minutiae extraction. Regarding record formats, due to the application of fingerprint biometrics to systems based on smart cards, compact record formats are also defined to cope with memory and transmission speed limitations of such devices.

Fingerprint scientists have defined more than 150 different types of minutiae [5]. Within this Standard, minutiae types are simplified to the following: (1) ridge ending, (2) ridge bifurcation, and (3) other. The location of each minutiae is determined by its horizontal and vertical position within the image. To determine such location a coordinate system is to be defined. Figure 3 shows how such coordinate system is chosen. Granularity to be taken to determine location is of one hundredth of a millimetre for the normal format, while just one tenth of a millimetre for card compact formats.

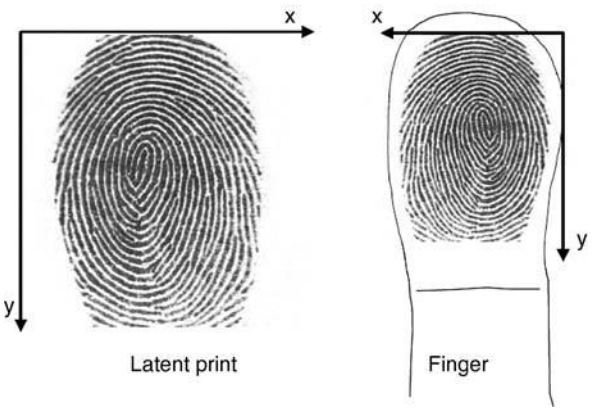
Figure 4 shows the different ways to consider the location of a minutiae. (1) represents a Ridge Ending, encoded as a Valley Skeleton Bifurcation Point, (2) shows how to locate a Ridge Bifurcation, encoded as a Ridge Skeleton Bifurcation Point, Finally (3) illustrates how to locate a Ridge Ending encoded as a Ridge Skeleton Endpoint. How to determine the encoding of ridge ending actually used in a specific dataset is a subject currently under revision in the standard. The other types of minutia have to be coded consistent with the Standards (see details in [3]).

To define the minutiae direction, its angle has to be determined. This Standards specifies that the angle is obtained, increasing counter-clockwise rotation starting from the horizontal axis to the right of the location of the minutiae point. The angle is encoded in a unsigned single byte, so the granularity is 1.40625° per bit (360/256). Figure 4 also illustrates how the angle is determined.

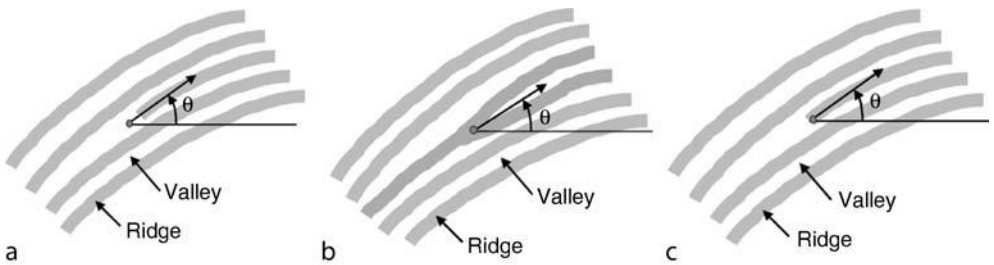
Additional information that may be included in a minutiae-based record are cores, deltas, and ridge crossings to neighboring minutiae.

With all these definitions, the two major format types defined by this International Standard are: (1) record format, and (2) card format. The structure of the record format is summarized in the following paragraphs and for additional details refer to the standard [3].

- A fixed-length (24-byte) record header containing information about the overallrecord, including the number of fingers represented and the overall record length in bytes:
  - Format identifier (4 bytes with the hexadecimal value 0x464D5200) and Version number (coded in another 4 bytes)
  - Record length (in bytes) including all finger images within that record (coded in 4 bytes)
  - Capture device ID (2 bytes)
  - Size of the image in pixels (2 bytes for X dimension, and 2 bytes for Y dimension)
  - Image resolution in pixels per centimetre (2 bytes for X and 2 bytes for Y)
  - Number of finger views included in the record
  - 1 byte reserved for future use



Finger Data Interchange Format, Standardization.
Figure 3 Coordinate System for Minutiae Location. Image taken from [3].



Finger Data Interchange Format, Standardization. **Figure 4** Illustration of location of minutia. Image taken from [3].

- A Single Finger record for each finger/view, consisting of:
  - A fixed-length (4-byte) header containing information about the data for asingle finger, including the number of minutiae:
    - Finger position (1 byte)
    - View number (4 bits) and Impression type (4 bits, to make a 1 byte in total)
    - Finger quality (1 byte)
    - Number of minutia (1 byte)
  - A series of fixed-length(6-byte) minutia descriptions:
    - Minutia type (2 bits) and X location in pixels (14 bits)
    - 2 bits reserved and Y location in pixels (14 bits)
    - Minutiae angle (1 byte)
    - Quality of minutiae (1 byte)
  - One or more "extended" data areas for each finger/view, containing optional or vendor-specific information. It starts always with 2 bytes which determine the length of Extended Data Block. If this is 0x0000, no Extended Data is included. If it has a nonnull value, then it is followed by vendor-specific data which could include information about ridge count, cores and deltas, or cell information.

Regarding the card formats, the current version of the standard allows 2 sub-formats: (1) normal format (also referred as 5-byte minutiae), and (2) compact format (also known as 3-byte minutiae). The way minutia are coded in each format is described below.

- Card normal format (like the record format, but removing quality information):
  - Minutia type (2 bits) and X location in pixels (14 bits)
  - 2 bits reserved and Y location in pixels (14 bits)
  - Minutiae angle (1 byte)
- Card compact format:
  - X coordinate (8 bits) considering a unit of 10^{-1}mm
  - Y coordinate (8 bits) considering a unit of 10^{-1}mm
  - Minutia type (2 bits) using the same coding as with the card normal format
  - Angle (6 bits) having a granularity of 360/64

Another important aspect related to card formats is that as they are intended to be used with devices with limited memory and processing power, the number of minutia may be restricted, and in such case, truncation is needed. Additionally in Match-on-Card systems, to reduce algorithm complexity, minutia may need to be sorted in a certain way. And finally, the way data is exchanged differs from the traditional CBEFF format. This International Standard covers all such cases. The reader is suggested to refer to the last version of the Standard [3] for further details.

The minutia standard is used e.g., by the ILO (International Labour Organization) in its seafarers identity card and in several national ID card implementations including Thailand and Spain [6].

Spectral Data of a Fingerprint

Part 3 of the 19794 series of standards deals with a format suitable to process fingerprints when using morphological approaches. But as seen in additional Fingerprint entries in this Encyclopedia, there are other approaches to perform biometric identification using fingerprints. Some of those approaches relate to the spectral information of the fingerprint. Algorithms using spectral data look at the global structure of a finger image rather than certain local point patterns. In such cases, 19794-2 is of no use and the only possibility would be to use the whole image as stated in 19794-4, which has the inconvenience of requiring the storage and/or transmission of a large amount of data. This could be inconvenient if not blocking for some applications.

In order to provide a new data format that could increase interoperability among spectral based solutions, reducing the amount of data to be used, 19794-3 has been developed under the title of "Information technology - Biometric data interchange formats - Part 3: Finger pattern spectral data" [7]. In fact, this International Standard deals with three major approaches in spectral based biometrics (wavelet based approaches are not supported by this standard).

1. Quantized co-sinusoidal triplets
2. Discrete Fourier transform
3. Gabor filters

After declaring the basic requirements for the original image in order to be considered for these algorithms (same coordinate system as in 19794-2, 255 levels of grey with 0 representing black and 255 being white, and

dark colours corresponding to ridges while light pixels corresponding to valleys), and describing all the above mentioned technologies, the Standards focuses on the record structure (for details refer to [7]), which is:

- A variable-length record header containing information about the overall record, including:
  - Format identifier (4 bytes with the hexadecimal value 0x46535000) and Version number (coded in another 4 bytes)
  - Record length (in bytes) including all fingers within that record (coded in 4 bytes)
  - Number of finger records included (1 byte)
  - Image resolution in pixels per centimetre (2 bytes for X direction and 2 bytes for Y direction)
  - Number of cells (2 bytes for X direction and 2 bytes for Y direction)
  - Number of pixels in cells (2 bytes for X direction and 2 bytes for Y direction)
  - Number of pixels between cells centres (2 bytes for X direction and 2 bytes for Y direction)
  - SCSM (Spectral component selection method - 1 byte), which can be 0, 1, or 2. Depending on the value of this field the following fields could refer to type of window, standard deviation, number of frequencies, frequencies, number of orientations and spectral components per cell, and bit-depths (propagation angle, wavelength, phase, and/or magnitude)
  - Bit-depth of quality score (1 byte)
  - Cell quality group granularity (1 byte)
  - 2 bytes reserved for future use
- A single finger record for each finger, consisting of:
  - A fixed-length (6-byte) header containing information about the data for asingle finger:
    Finger location (1 byte)
    Impression type (1 byte)
    Number of views in single finger record (1 byte)
    Finger pattern quality (1 byte)
    Length of finger pattern spectral data block (2 bytes)
  - A finger pattern spectral data block:
    View number (1 byte)
    Finger pattern spectral data
    Cell quality data
  - An extended data block containing vendor-specific data, composed of block length (2 bytes), area type code (2 bytes), area length, and area.

As in 19794-2, this International Standard also defines the Data Objects to be included for a card format, with the reduction in granularity recommended (for further details see [7]).

Some of the leading fingerprint verification algorithms rely on spectral data or a combination of spectral data and minutiae. This standard could enhance the interoperability and performance of large scale identification systems such as criminal or civil Automatic Fingerprint Identification Systems (AFIS).

Skeletal Data of a Fingerprint

Finally 19794-8 titled "Information technology - Biometric data interchange formats - Part 8: Finger pattern skeletal data" [8] deals with the format for representing fingerprint images by a skeleton with ridges represented by a sequence of lines. Skeletonization is a standard procedure in image processing and generates a single pixel wide skeleton of a binary image. Moreover the start and endpoints of the skeleton ridge lines are included as real or virtual minutiae, and the line from start to endpoint is encoded by successive direction changes.

For minutiae location and coding, much of the 19794-2 card format is used, but here the position of a ridge bifurcation minutiae shall be defined as the point of forking of the skeleton of the ridge. In other words, the point where three or more ridges intersect is the location of the minutia. No valley representation is accepted under this International Standard. Another difference with 19794-2 card formats, is that in this Standard no other-type minutiae is considered (if a minutiae has more than three arms, like a trifurcation, it is considered a bifurcation), and that along this standard codes for "virtual minutiae" are used.

Skeleton lines are coded as polygons. Every line starts with a minutiae, and it is followed by a chain of direction changes (coded with the granularity stated in the record header), until it reaches the final minutiae. Several rules are defined in the standard (see [8] for further reference).

All that information is coded in a record with the following structure (limiting values as well as recommended values can be found in [8]):

- A fixed-length (24-byte) record header containing:
  - Format identifier (4 bytes with the hexadecimal value 0x46534B00) and Version number (coded in another 4 bytes)

  - Record length (in bytes) including all finger images within that record (coded in 4 bytes)
  - Capture device ID (2 bytes)
  - Number of finger views in record (1 byte)
  - Resolution of finger pattern in pixels per centimetre (1 byte)
  - Bit depth of direction code start and stop point coordinates (1 byte)
  - Bit depth of direction code start and stop direction (1 byte)
  - Bit depth of direction in direction code (1 byte)
  - Step size of direction code (1 byte)
  - Relative perpendicular step size (1 byte)
  - Number of directions on 180° (1 byte)
  - 2 bytes reserved for future use
- A single finger record for each finger/view, consisting of:
  - A fixed-length (10 bytes) header:
    - View number (1 byte)
    - Finger position (1 byte)
    - Impression type (1 byte)
    - Finger quality (1 byte)
    - Skeleton image size in pixels (2 bytes for X-direction, 2 bytes for Y-direction)
    - Length of finger pattern skeletal data block (2 bytes)
  - The variable length fingerprint pattern skeletal description:
    - Length of finger pattern skeletal data (2 bytes)
    - Finger pattern skeletal data
    - Length of skeleton line neighbourhood index data (2 bytes)
    - Skeleton line neighbourhood index data
  - An extended data block containing the extended data block length and zero or more extended data areas for each finger/view, defining length (2 bytes), area type code (2 bytes), area length (2 bytes), and data.

This International Standard also defines two card formats, a normal one and a compact one. As with other parts, this means more limiting constraints to code data tighter and the definition of the Data Objects needed (for details refer to [8]).

The skeleton format is used in scientific research [9] and by vendors, implementing Match-on-Card.

Further Steps

The fingerprint parts of ISO 19794 were published as International Standards in 2005 and 2006. All the parts are currently under revision. A major task in the revision process is to address some defects and include a common header format for all the parts. Some references and vocabulary are needed to be updated to harmonize the relation of these standards within the ISO standardization landscape. The finger minutia standard ISO 19794-2 is probably the most prominent format in this series and is most frequently used by industry, government, and science. Interoperability tests have shown that the current standard allows some room for interpretation. This will be compensated by an amendment to describe the location, orientation, and type in more detail. Another aspect in the current revision of the standard is to reduce the number of format types from currently ten to a maximum of two. Experts from all continents and various backgrounds meet on a regular basis to lay down the future of the standards. The delegates take care of current requirements in terms of technology and applications.

Summary

To provide interoperability in storing and transmitting finger-related biometric information, four standards are already developed to define the formats needed for raw images, minutia-based feature vectors, spectral information, and skeletal representation of a fingerprint. Beyond that, other standards deal with conformance and quality control, as well as interfaces or performance evaluation and reporting (see related entries below for further information).

Related Entries

► Biometric Data Interchange Format
► Common Biometric Exchange Framework Formats
► Conformance Testing for Biometric Data Interchange Formats, Standardization of
► Fingerprint Recognition
► International Standardization of Biometrics

References

1. ISO/IEC: 19794-4:2005 - information technology - biometric data interchange formats - part 4: Finger image data (2005)
2. ISO/IEC: 19785-1:2005 - information technology - common biometric exchange formats framework - part 1: Data element specification (2005)
3. ISO/IEC: 19794-2:2005 - information technology - biometric data interchange formats - part 2: Finger minutiae data (2005)
4. Galton, F.: Finger Prints. Macmillan, London (Reprint: Da Capo, New York, 1965) (1892)
5. Moenssens, A.: Fingerprint Techniques. Chilton Book Company, London (1971)
6. Spanish-Homeland-Ministry: Spanish national electronic identity card information portal (in spanish). http://www.dnielectronico.es/ (2007)
7. ISO/IEC: 19794-3:2006 - information technology - biometric data interchange formats - part 3: Finger pattern spectral data (2006)
8. ISO/IEC: 19794-8:2006 - information technology - biometric data interchange formats - part 8: Finger pattern skeletal data (2006)
9. Robert Mueller, U.M.: Decision level fusion in standardized fingerprint match-on-card. In: 1-4244-0342-1/06, ICARCV 2006, Hanoi, Vietnam (2006)

Finger Geometry, 3D

Sotiris Malassiotis
Informatics and Telematics Institute, Center for Research and Technology Hellas, Thessaloniki, Greece

Synonym

3D hand biometrics

Definition

Biometrics based on 3D finger geometry exploit discriminatory information provided by the 3D structure of the hand, and more specifically the fingers, as captured by a 3D sensor. The advantages of current 3D finger biometrics over traditional 2D hand geometry authentication techniques are improved accuracy, the ability to work in contact free mode, and the ability to combine with 3D face recognition using the same sensor.

Introduction

The motivation behind 3D finger geometry biometrics is the same as with 3D face recognition. The 3D geometry of the hand as captured by a 3D sensor offers additional discriminatory information while being invariant to variations such as illumination or pigment of the skin, compared with an image captured with a plain 2D camera. The current accuracy and resolution of 3D sensors are not adequate for capturing fine details on the surface of the fingers such as skin wrinkles over the knuckles, but is sufficient to measure, local curvature, finger circumference, or finger length.

Another motivation comes from a limitation of current hand geometry recognition systems, that is obtrusiveness. The user is required to put his/her hand on a special platter with knobs or pegs that constrain the placement of the hand on the platter. This step greatly facilitates the process of feature extraction by guaranteeing a uniform background and hand posture. Thus it guarantees very good performance. However, several users would find touching of the platter unhygienic, while others would face difficulty correctly placing their hands (for example children or older people with arthritis problems). Since 3D data can facilitate the detection of the hand and fingers, even in a cluttered scene, the above constraint may be raised and the biometric system becomes more user friendly.

Since the placement of the hand is not a constraint, one may then combine 3D finger geometry with 3D face using the same 3D sensor. The user either places his/her hand on the side of the face or in front of the face. In the first case, face and hand biometric features are extracted in parallel, while in the second case sequentially and the scores obtained are finally combined. This combination has demonstrated very high accuracy even under difficult conditions.

State-of-the-Art

3D hand geometry biometrics is a very recent research topic and therefore, only a few results are currently available.

The first to investigate 3D geometry of the fingers as a biometric modality were Woodard and Flynn [1].

They used a 3D laser scanner to capture range images and associated color images of the back of the hand. The users were instructed to place their palm flat against a wall with uniform color and remove any rings. For each subject out of 132, four images were captured in two recording sessions one week apart. An additional session was also performed a few months later with 86 of the original subjects and 89 new subjects.

The authors used the color images to perform segmentation of the hand from the background. A combination of skin-color detection and edge detection was used. The resulting hand segmentation is used to extract the hand silhouette from which the boundaries of index, middle, and ring fingers are detected. Then for each detected finger a mask is constructed and an associated normalized (with respect to pose) range image is created.

For each valid pixel of the finger mask in the output image, a ▶ surface curvature estimate is computed with the corresponding range data. The principal curvatures are estimated first by locally fitting a bicubic Monge patch on the range data to deal with the noise in the data. However, the number of pixels in the neighborhood of each point that are used to fit the patch has to be carefully selected, otherwise fine detail on the surface may be lost. The principal curvatures are subsequently used to compute a shape index, which is a single measure of curvature.

The similarity between two finger surfaces may be computed by estimating the normalized correlation coefficient among the associated shape index images. The average of the similarity scores obtained by the three fingers demonstrated the best results when used for classification.

Recognition experiments demonstrated an 95% accuracy, falling to 85% in the case that probe and gallery images are recorded more than one week apart. This performance was similar with that reported by a 2D face recognition experiment. The authors managed to cope with this decline in performance due to time lapse by matching multiple probe images with multiple gallery images of the same subject. Similarly, the equal error rate obtained in verification experiments, is about 9% when a single probe image is matched against a single gallery image and falls to 5.5% when multiple probe and gallery images are matched.

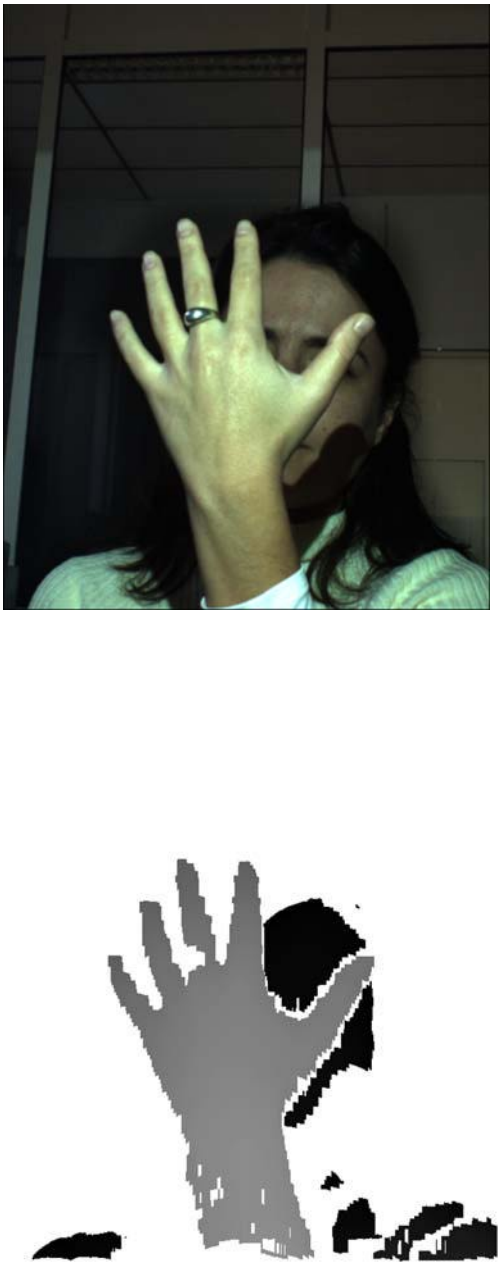
The above results validated the assumption that 3D finger geometry offers discriminatory information and may provide an alternative to 2D hand geometry recognition. However it remains unclear how such an approach will fair against a 2D hand geometry based system, given the high cost of 3D sensor.

"The main advantage of a biometric system based on 3D finger geometry is its ability to work in an unobtrusive (contact-free) manner [2]." They propose a biometric authentication scenario where the user freely places his hand in front of his face with the back of the hand visible from the 3D sensor. Although the palm should be open with the fingers extended, small finger bending and moderate rotation of the hand plane with respect to the camera are allowed as well as wearing of rings.

The acquisition of range images and quasi-synchronous color images are achieved using a real-time 3D sensor, which is based on the structured light approach. Thus, data are more noisy and contain more artifacts compared with those obtained with high-end laser scanners. Using this setup, the authors acquired several images of 73 subjects in two recording sessions. For each subject, images depicting several variations in the geometry of the hand were captured. These included, bending of the fingers, rotation of the hand, and presence or absence of rings (see fig. 1).

The proposed algorithm starts by segmenting the hand from the face and torso using thresholding and subsequently from the arm using an iterative clustering technique. Then, the approximate center of the palm and the orientation of the hand is detected from the hand segmentation mask. These are used to locate the fingers. Homocentric circular arcs are drawn around the center of the palm with increasing radius excluding the lower part of the circle that corresponds to the wrist. Intersection of these arcs with the hand mask gives raise to candidates of finger segments, which are then clustered to form finger bounding polygons. This approach avoids using the hand silhouette, which is usually noisy and may contain discontinuities, e.g., in the presence of rings. The initial polygon delineating each finger is refined by exploiting the associated color image edges.

Then, for each finger two signature functions are defined, parameterized by the 3D distance from the finger tip computed along the ridge of each finger and measuring cross-sectional features. Computing features along cross-sections offers quasi- invariance to bending. The first function corresponds to the width of the finger in 3D, while the second corresponds to the mean curvature of the curve that is defined by the 3D points corresponding to the cross-section at the specific



Finger Geometry, 3D. Figure 1 (**a**) Color and (**b**) range image in the hand geometry acquisition setup of [2].

point. Twelve samples are uniformly computed from each signature function and each finger giving rise to 96 measurements (the thumb was excluded) that are used for classification. Matching between hand geometry of probe and gallery images is estimated as the L_1 distance between the associated measurement vectors.

Experimental results are similar with those reported in [1]. Rank-1 identification rates range from 86 to 98% depending on using a single or multiple probe images of the same subject respectively. Corresponding equal error rates are 5.8 and 3.5%. The benefit of the approach in [2] is that the algorithm can withstand moderate variations in hand geometry thus allowing for contact free operation.

Malassiotis et. al [2] conclude that given the current results biometric systems that exploit 3D hand geometry would be more suitable in low security applications such as personalization of services and attendance control where user-friendliness is prioritized over accuracy. However, there is another possible application in systems combining several biometrics. In particular, the combination of 3D face modality with 3D finger geometry was shown to offer both high accuracy and also be relatively unobtrusive.

Woodard et al. [3] compared the recognition performance of 3D face, 3D ear, and 3D finger surface as well as their combination. The original 93% obtained using 3D face geometry was improved to 97% when this was combined with the other two modalities.

Tsalakanidou et al. [4] also combined 2D+3D face recognition with 3D finger geometry recognition, in the presence of several variations in shape and appearance of the face and hand. According to their application scenario, the 3D sensor grabs first images of the user's face and then the user is asked to place his hand in front of his face and another set of images is acquired. The scores obtained using facial and hand features respectively are normalized and fused to provide a single score on which identification/verification is based. An Equal Error Rate equal to 0.82% and a rank-1 identification rate equal to 100% was reported for a test-set comprised of 17,285 pairs of face and hand images of 50 subjects depicting significant variations.

The above results validate our original claim that 3D face geometry + 3D finger geometry may provide both high accuracy and user acceptance while sharing the same sensor for data acquisition.

Challenges and Prospects

Biometric authentication/identification using 3D finger geometry is a very recent addition in the compendium of 3D biometrics. Although the potential of this technique has been already demonstrated, several research challenges have to be addressed before commercial applications using this modality emerge.

Performance of techniques based on 3D finger geometry depends much more on the quality of range data than 3D face recognition. Although, some of the fine detail on the finger surface may be captured using high-end (and therefore very expensive) 3D scanners, this is not the case with low-cost systems. Such detail (e.g. the wrinkles of the skin) may be alternatively detected if associated brightness images are used. In this case, 3D information may be used to facilitate the localization of the finger and knuckles and 2D images may be subsequently used to extract the skin folding patterns. Also, both studies in the literature do not use the thumb finger, which however, seems to exhibit larger variability from subject to subject than the rest of the fingers.

Further research is also needed to address the problem of the variability in the shape and appearance depicted on the hand images. Future techniques should be able to deal with significant finger bending, partial finger occlusion, and rotation of the hand with respect to the camera and also be generic enough to cope with different hand sizes and deformed finger due to accident or aging.

In summary, 3D finger biometrics retain the benefits of traditional 2D hand geometry biometrics especially with respect to privacy preservation, while demonstrating similar or better performance. In addition, 3D finger biometrics may be applied with less strict constraints on the placement of the hand and the environment, which makes them suitable for a larger range of low to medium security applications.

Since correlation of finger geometry features with other discriminative features of the human body is known to be very low, 3D finger geometry may be efficiently combined with other biometrics in a multimodal system. In this case, this technology may be applied to high security scenarios.

Related Entries

► 3D-Based Face Recognition
► Hand Geometry

References

1. Woodard, D.L., Flynn, P.J.: Finger surface as a biometric identifier. Comput. Vision Image Understand **100**, 357–384 (2005)
2. Malassiotis, S., Aifanti, N., Strintzis, M.G.: Personal Authentication Using 3-D Finger Geometry. IEEE Trans. Inform. Forens. Secur. **1**(1), 12–21 (2006)
3. Woodard, D.L., Faltemier, T.C., Yan, P., Flynn, P.J., Bowyer, K.W.: A Comparison of 3D Biometric Modalities. In: Proc. Comput. Vision Pattern Recogn. Workshop, pp. 57–60 (2006)
4. Tsalakanidou, F., Malassiotis, S., Strintzis, M.G.: A 3D Face and Hand Biometric System for Robust User-Friendly Authentication. Pattern Recogn. Lett. **28**(16), 2238–2249 (2007)

Finger Pattern Spectral Data

Set of spectral components derived from a fingerprint image that may be processed (e.g., by cropping and/or down-sampling).

► Finger Data Interchange Format, Standardization

Finger Vein

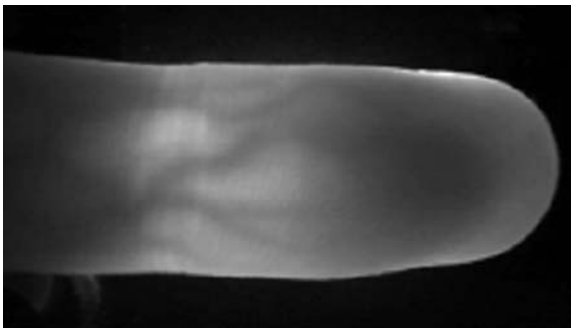
Hisao Ogata Mitsutoshi Himaga
Hitachi-Omron Terminal Solutions, Corp.
Owari-asahi City, Aichi, Japan

Definition

Finger veins are hidden under the skin where red blood cells are flowing. In biometrics, the term vein does not entirely correspond to the terminology of medical science. Its network patterns are used for authenticating the identity of a person, in which the approximately 0.3–1.0 mm thick vein is visible by ► near infrared rays. In this definition, the term finger includes not only index, middle, ring, and little fingers, but also the thumb.

Introduction

Blood vessels are not exposed and their network patterns are normally impossible to see without the range of visible light wavelength. The approximately 0.3–1.0 mm vein which constitutes the network patterns are visualized by near infrared rays. Figure 1 shows a visualized finger vein pattern image. It is well known that hemoglobin absorbs near infrared rays more than other substances that comprise the human



Finger Vein. **Figure 1** Extracted finger vein image.

body. Since most of the hemoglobin of human body exists in red blood cells that are flowing inside blood vessels, the blood vessel network patterns can be seen as a dark area by infrared imaging systems. Vascular network patterns inside finger of an individual are visualized by utilizing this optical characteristic of hemoglobin. Therefore the network patterns can be used as a biometric modality by appropriate imaging technologies. As the diameters of arteries are as small as approximately 1/3 of those of targeted veins in finger, it is reasonable to assume that most of the visualized blood vessels are veins. This is why many of vascular biometric technologies are known as "vein" biometrics, though arteries and veins are equally visualized by infrared light and normally treated in the same manner.

Kono et al. developed a near-infrared finger vein reader prototype and demonstrated its effectiveness in 2000 [1], and further evaluated the performance of the proposed biometric modality by using sample data collected from 678 subjects and reported very positive results in 2002 [2].

There are two major approaches to visualize vascular patterns for biometric use, namely the light penetration method and the light reflection method. The light penetration method utilizes the infrared light transmitted through the target object, while the light reflection method makes use of the light reflected by the target. The light reflection method is not usually the first choice unless it is necessary because it is difficult to handle the reflected images that may contain saturated (over-exposed) areas or texture on the skin surface. The contrast of the images captured by penetrating light is generally higher than that captured by reflected light. The high contrast images result in high accuracy of authentication because more information to distinguish the network patterns can be extracted from the high signal to noise ratio image. However, the light reflection method is only a choice in case of imaging thick target objects such as palm vein or the back-of-the-hand vein in which near infrared rays are not transmitted through the body. Fingers are only parts of a human body which can be easily presented to an authentication device, and from which clear pattern images can be captured by using "light penetration method." Therefore, finger vein biometrics is recognized as one of the most reliable and stable biometric modalities.

Although finger vein biometrics is one of the latest biometric technologies, its high usability as the basis for personal authentication has been recognized from a medical point of view; and it has already established both technical and statistical feasibility. In the following sections, medical opinions describe how the finger vein conforms to three desirable properties for biometrics. The uniqueness of Finger Vein was also evaluated in statistical approach.

Medical Opinions Concerning Finger Vein Authentication Technology

In 2006, Central Research Laboratory, Hitachi, Ltd. (Tokyo, Japan) [3] and Hitachi-Omron Terminal Solutions, Corp. (Tokyo, Japan) [4] held a series of four Finger Vein Authentication Workshops, which was attended by representative Japanese researchers. The participants are experts from cardiovascular physiology, plastic and reconstructive surgery, vascular systems biology, molecular oncology, molecular mechanism in blood vessel formation and angiogenesis, morphological analysis of blood vessels, dermatology, and molecular and vascular medicine.

Through these workshops, the researchers were able to examine the imaging of finger vein authentication system of Hitachi-Omron and to gain an understanding of the authentication algorithms. The workshops were an opportunity to obtain from researchers several improvement medical opinions concerning finger vein authentication technology that are set forth below.

a) Universality

Veins and arteries are essential for circulating oxygen and nutrients to the finger tissues, and it is

a fact known to medical science that the approximately 0.3–1.0 mm thick vein in the skin surface layer that is targeted for the authentication basically exists in all people.

b) Uniqueness

In ontogenesis, the patterning of the vascular network undergoes change from its initial state, and the arteriovenous network is formed subject to the effects of low oxygen and blood flow. This process takes place under genetic constraints, but is not deterministic; it includes many probabilistic elements. Thus, there will be large individual differences in the pattern of the vein that is used for authentication, and its utility as the basis for personal authentication will be high.

c) Permanence

The basic pattern of the blood vessels is formed during the fetal stage. Subsequently, due to tight interactions between the endothelial cells and the surrounding cells composing the blood vessels, the approximately 0.3–1.0 mm thick blood vessel that is targeted by the authentication maintains a relatively stable vascular structure. In addition, the blood vessel targeted by the authentication is assured of a permanent flow of blood, and in healthy adults it is extremely unlikely to be lost with aging. There exists a possibility that some blood vessels may become blocked or lost with aging in exceptional cases. Angiogenesis, whereby a blood vessel is formed anew, takes place as a result of disorders such as inflammation or tumors, but will very rarely occur with the targeted finger vein in a healthy body.

d) Racial/ethnic differences

No large racial or ethnic variations are known in the patterns relevant for personal identification.

Uniqueness in Statistical Approach

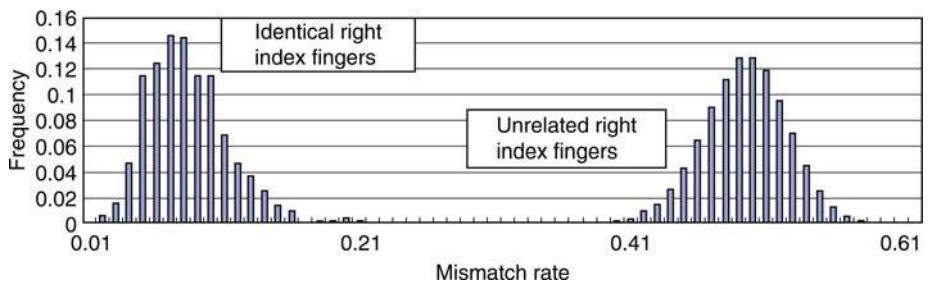
In 2007, Yanagawa et al. demonstrated the diversity of human finger vein patterns by conducting statistical analysis based on sample data collected from 506 subjects. They also proved the feasibility(reliability) of using finger vein patterns for personal identification by evaluating false acceptance rates (FAR) and false rejection rates (FRR) based on mathematical models [5].

a) Diversity of finger vein patterns

Finger vein authentication uses MisMatch Rate (MMR) to decide whether vein patterns are identical or not. MMR is defined as

$$\text{MMR} = \frac{\text{total number of mismatched pairs}}{\text{total number of pixels classified into vein in the two finger patterns}}$$

Figure 2 shows histograms of the MMR computed from 1,012 (= 506 person × 2) pairs of identical right index fingers and 255,530 (= 506 × 505) pairs of unrelated right index fingers. The figure shows that two histograms are separated, indicating the significant difference of vein patterns of the right index finger between individuals. The histograms of MMR derived from the pairs of unrelated right index fingers are almost overlapped with other pair combinations; a right index finger and a right middle finger of an identical person, a right index finger and a left index finger of an identical person. These observations indicate that two fingers are identical if and only if they are the same finger in the same hand of the same person, and all the other cases can be treated simply as unrelated.



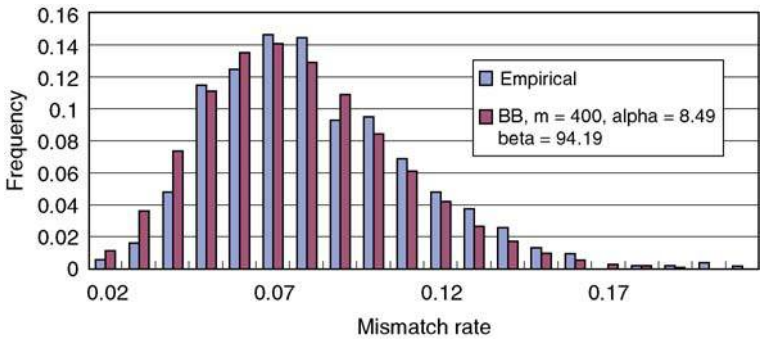
Finger Vein. Figure 2 Histograms of mismatch rates computed (MMR) from 1,012 pairs of identical right index finger and 255,530 pairs of unrelated index finger.

b) Reliability estimation of personal identification by mathematical models

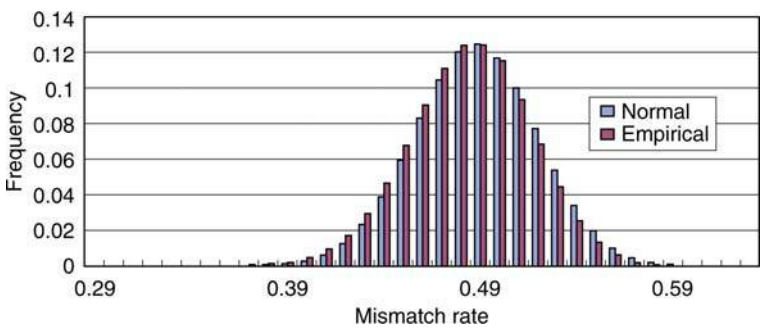
The validity of our personal identification is evaluated by two probabilities inherent to the device, the FRR and the FAR. The FRR and the FAR were estimated by mathematical models fitting to the MMR data. Figure 3 shows the histograms of MMR computed from identical right index fingers (empirical 1,012 pairs) and fitted beta-binominal distribution, demonstrating the fitting is fairly good. Figure 4 shows the histogram from 2,540,120 unrelated pairs (empirical). The histogram and the normal distribution N(0.4859, 0.03082) shows pretty good correspondence. Table 1 shows the estimated FRR and FAR from the beta-binominal distribution and the normal distribution respectively for selected values of the cut-off points. For example, the FRR is 3.16E-6 and the FAR is 1.31E-12 at the cut-off point of 0.270 on the table while the FRR is 1.0E-4 and the FAR is 1.0E-6 in the official accuracy specification of actual authentication products. Accordingly, finger vein pattern itself has potential to achieve quite high accuracy.

Summary

Today, finger vein biometrics is recognized as one of the most reliable and secure biometric modalities and is applied to a variety of security systems. As described here, it has already established both of statistical feasibility and its high usability as the basis for personal authentication is recognized from the point of view of the medical opinions. The FRR derived from the mathematical models fitted to the empirical histograms is 1.31E-12, while the official FAR of the current finger vein authentication products is 1.0E-6. Accordingly, finger vein pattern itself has the potential to achieve quite high accuracy. Unlike conventional biometric features such as finger print, vascular network patterns cannot be observed without using specially designed



Finger Vein. **Figure 3** Histograms of mismatch rates (MMR) computed from 1,012 pairs of identical right index finger (empirical) and Beta-Binomial distribution with m = 400, α = 8.49 and β = 94.19.



Finger Vein. **Figure 4** Histograms of mismatch rates (MMR) computed from 2,540,120 unrelated pairs of right index finger (empirical) and normal distribution with mean = 0.4859 and s.d. = 0.0308.

Finger Vein. **Table 1** Estimated false acceptance rate (FRR) and false rejection rate (FAR)

Cut-off point	FRR	FAR	95% c.i.	Of FAR
0.270	3.16E-06	1.31E-12	6.32E-13	2.56E-12
0.275	2.03E-06	4.10E-12	2.07E-12	7.80 E-12
0.280	1.30E-06	1.25E-11	6.41E-12	2.45 E-11
2.285	8.23E-07	3.73E-11	2.00E-11	6.96 E-11
2.290	5.20E-07	1.08E-10	5.82E-11	1.94 E-10
2.295	3.27E-07	3.07E-10	1.74E-10	5.49 E-10
0.300	2.04E-07	8.47E-10	4.84 E10	1.46 E-09
0.305	1.27E-07	2.28E-09	1.35E-10	3.85 E-09
3.310	7.86 E-08	5.97E-09	3.69E-11	9.81 E-09

Finger Vein. **Figure 5** ATM equipped with a finger vein reader (Courtesy of Hitachi-Omron Terminal Solutions, Corp.).

equipment, and thus it is extremely difficult to steal or duplicate the biometric information. Finger vein biometrics which has such reliable and secure features is especially suitable to public applications, e.g., banking systems, medical systems, and passport controls. Its banking applications (Fig. 5) remain one of the largest and the most successful set of applications for this state-of-the-art biometric modality; and it is anticipated that more than a quarter of ATMs in Japan will be equipped with finger vein readers by the end of 2008.

Related Entries

- ▶ Finger Vein Feature Extraction
- ▶ Finger Vein Imaging
- ▶ Finger Vein reader
- ▶ Hand Veins

References

1. Kono, M., Ueki, H., Umemura,S.: A new method for the identification of individuals by using vein pattern matching of a finger. In: Proceedings of the Fifth Symposium on Pattern Measurement (Yamaguchi, Japan), pp. 9–12 (2000) (in Japanese)
2. Kono, M., Ueki, H., Umemura, S.: Near-infrared finger vein patterns for personal identification. Appl. Opt. **41**(35), 7429–7436 (2002)
3. Hitachi Central Research Laboratory, http://www.hitachi.comfrd/cr//
4. Hitachi-Omron Terminal Solutions, Corp., http://www.hitachi-omron-ts.com/index.html
5. Yanagawa, T., Aoki, S., Ohyama, T.: Human finger vein images are diverse and its patterns are useful for personal identification. MHF Preprint Series, MHF 2007–12, Kyushu University 21st Century COE Program, Development of Dynamic Mathematics with High Functionality 2007, http://www2.math.kyushu-u.ac.jp/coe/report/pdf/2007-12.pdf
6. Jain, A.K., Bolle, R., Pankanti, S.: Biometrics: Personal Identification in Networked Society. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999

Finger Vein Authentication Device

► Finger Vein Reader

Finger Vein Biometric Algorithm

Mitsutoshi Himaga
Hitachi-Omron Terminal Solutions, Corp. Tokyo, Japan

Synonym

Finger vein feature segmentation

Definition

Finger vein biometric algorithm is a series of software processes to authenticate a person by using biometric features extracted from his or her finger vein patterns. The algorithm is typically comprised of two major processes, namely, a finger vein feature extraction part and a matching algorithm part.

Introduction

Finger vein feature extraction, along with finger vein imaging technology, is a core technology in finger vein authentication. By applying this process, a simple ► *raw finger vein image* is converted into meaningful biometric data that can be used to identify a person. The finger vein feature extraction is executed in both the enrollment process and the authentication process of a finger vein biometric system. In the enrollment process, the extracted biometric data is used to create template data together with the associated personal information such as username or identification numbers. In the authentication process, the finger vein feature extraction is applied to each frame of the scanned image prior to the matching process with the pre-registered template data.

The selection of the biometric features is dependent on the extraction algorithm, and therefore, the features extracted by one algorithm can be very different from those extracted by another, even for an identical finger. This means that a template produced by one finger vein system may not be compatible with another. There are multiple manufacturers who have commercialized finger vein authentication systems; however, the compatibility of finger vein templates is not guaranteed in many cases.

Requirements for the Finger Vein Biometric Algorithm

Unlike other biometrics such as finger print, finger vein patterns do not leave any trace and can only be observed by using a purpose-made imaging device. This makes it extremely difficult to steal or duplicate the biometric features, which comprises one of the many reasons to use this biometric modality. On the other hand, from technical point of view finger vein biometrics requires some special image processing technology that enables the system to extract clear and stable biometric features. Since the quality of raw images of intra-body structure is generally very poor, a sophisticated illumination control and image processing technology is required. In other words, quite of lot of technical know-how is necessary to extract high quality biometric features from such low quality images that have a large individual variation. Considering the variety of the know-how, it is quite reasonable to assume that there are many implementations of finger vein biometric algorithms. The compatibility of the biometric information (i.e., templates) is, however, largely dependent on the biometric algorithm and, therefore, it is very important to design the algorithm so that the template can be widely applicable to a variety of applications.

Finger Vein Feature Extraction

As described in the previous section, the details of the finger vein feature extraction are not publicly available due to its secure nature as of the time of writing. However, there are a few technical papers reported by the leading manufacturer, Hitachi, Ltd. (Tokyo, Japan) [1]. One of the earliest finger vein feature extraction algorithms developed by the Central Research Laboratory (CRL) of Hitachi, Ltd. is briefly introduced in the below section [2].

The finger vein feature extraction process is as follows.

Step 1: Set a starting point.

An initial point is set at random within the area inside the finger.

Step 2: Set a group of candidate pixels for the next point.

A group of candidate pixels are selected from the neighborhood of the initial point by using a weighted random number. Considering the blood vessel paths, the weighting coefficients are configured by experiment so that horizontally connected pixels are more likely to be selected than vertically or diagonally connected pixels.

Step 3: Find the darkest path

All candidate pixels selected in Step 2 are tested to find the darkest direction. Each candidate point is evaluated by analyzing the intensity difference between the brightest pixel and the darkest pixel along the intensity profile orthogonally crossing to the vector made by the current pixel and the candidate pixel.

Step 4: Update the score

If the selected candidate pixel in Step3 has never visited during the current pass, the score of the candidate point is increased and the current point is moved to the candidate pixel. If the selected candidate pixel has ever visited or no pixel was selected in Step 3, Step 6 can be used directly.

Step 5: Go back to Step 2

Step 6: Repeat Step 1–4 for 3,000 times.

After repeating this process for 3,000 times, a map of the scores is created. As the above-mentioned algorithm traces the bottom of the intensity profile, or in other words, the darkest part within the area of the finger vein network, highly-scored pixels tend to be found in the middle of the blood vessels. Figure 1 shows the score map created by this algorithm. The score is normalized by the factor of 255 so that the map can be interpreted as an 8-bit greyscale image.

CRL introduced another finger vein feature extraction algorithm in 2002 [3], which is very different from the above algorithm.

Matching Algorithm

The matching algorithm for finger vein biometrics can also be implemented in many ways. A matching algorithm evaluated by Yanagawa et al. [4] is briefly described below as an example.

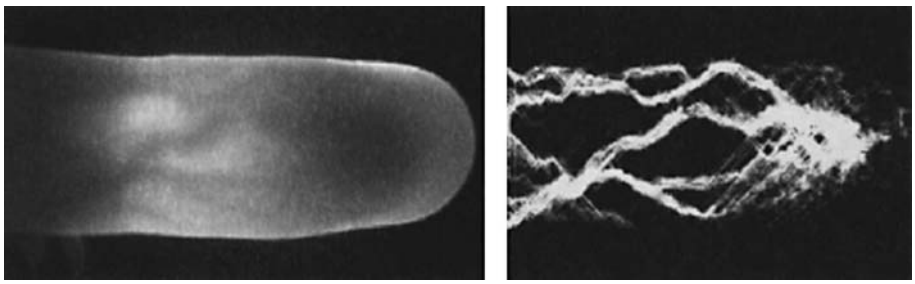
Yanagawa et al. published one of the very few technical papers in 2007 that describe a method to evaluate the similarity between the two finger vein patterns, in which they proved the feasibility of using finger vein patterns as biometric features from a statistical point of view. The similarity index they used for the statistical evaluation is as follows.

Pixels that consist of an extracted vein pattern are classified into three categories, namely, VEIN, AMBIGUOUS, and BACKGROUND. A pair of finger vein patterns to be evaluated is overlapped and compared pixel-by-pixel. If a pixel belongs to VEIN in the first pattern corresponds to a pixel belongs to BACKGROUND in the second pattern, the pair of pixels is regarded to be mismatched.

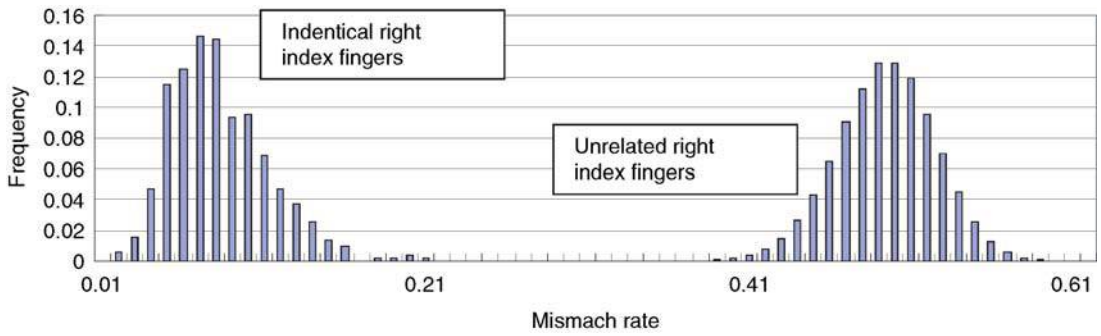
The mismatch rate (MMR) is defined as:

$$\mathrm{MMR} = \frac{\textit{The total number of mismatched pairs}}{\textit{The total number of pixels classified into VEIN in the two finger vein patterns}}$$

It is noted that MMR is not a symmetric index. Since pixels belonging to AMBIGUOUS and BACKGROUND in the first pattern are excluded from the calculation, the number of mismatched pairs varies depending on which pattern is regarded as the first pattern. Suppose a pair of finger vein patterns, R and L, have



Finger Vein Biometric Algorithm. **Figure 1** Visualised finger vein network (left) and its segmented pattern (right).



Finger Vein Biometric Algorithm. **Figure 2** Histograms of mismatch rates computed based on the right index figures.

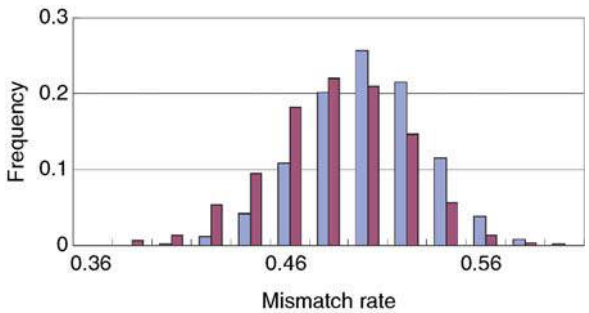
three corresponding pixels that are classified into AMBIGUOUS, VEIN, BACKGROUND and AMBIGUOUS, VEIN, VEIN, respectively, then there are no mismatched pairs when R is selected as the first image, while one mismatched pair is counted when L is select as the first image.

In order to evaluate the feasibility of using finger vein as a biometric feature, Yanagawa et al. collected finger vein patterns from 506 subjects (405 males and 101 females). They obtained multiple instances of index and middle finger vein pattern from each subject and compared them with the MMR value distributions of identical and non-identical vein pattern pairs.

Figure 2 shows the histogram of the MMR values calculated from the 1,012 pairs of identical right index fingers (i.e., 506 subjects $\times$ 2) and 255,530 unrelated pairs of right index fingers (i.e., 506 $\times$ 505). The figure shows that peaks of the two histograms are clearly separated, which indicates the significant inter-subject difference of finger vein patterns.

Figure 3 shows the histograms of MMR computed from 255,530 pairs of right middle fingers and right index fingers from identical person (dark bars) and 255,530 pairs of unrelated right index fingers (bright bars). The figure shows that two histograms are almost overlapped, indicating that the intra-subject differences of finger vein pattern are not significantly larger than inter-subject differences.

Table 1 shows the performance of MMR-based finger vein biometrics in terms of FAR and FRR. Yanagawa et al. estimated the FAR and the FRR based on the fitted normal distribution and the fitted beta-binominal distribution, respectively. They successfully demonstrated the supreme characteristics of biometrics by illustrating the two indices over several cut-off points



Finger Vein Biometric Algorithm. **Figure 3** Histograms of mismatch rates computed based on the right index fingers and middle fingers of identical person (dark) and those of unrelated individuals (bright).

(threshold values) together with 95% confidence intervals of the FAR. The figures in Table 1 are particularly better than the publicly announced FAR and FRR values of commercial products; at the cut-off point of 0.270 for instance, the statistical analysis indicates that the estimated FAR and FRR are as low as 1.31E-12 and 3.16E-6, respectively. These figures are far lower than the claimed FAR (1.0E-6) and FRR (1.0E-4) of commercial products, which implies that the biometric feature as such has a very preferable characteristic that can potentially achieve even higher accuracy.

These results strongly support the feasibility of finger vein biometrics and imply that indices such as MMR can effectively distinguish genuine patterns from others by applying an appropriate threshold value. The index described here is, however, quoted solely for the purpose of explanation, and therefore, it does not really represent the actual finger vein matching algorithm employed by commercial products.

Finger Vein Biometric Algorithm. **Table 1** Estimated false acceptance rate (FAR) and false rejection rate (FRR)

Cut-off point	FRR	FAR	95% c.i. of FAR	
0.270	3.16E-06	1.31E-12	6.32E-13	2.56E-12
0.275	2.03E-06	4.10E-12	2.07E-12	7.80E-12
0.280	1.30E-06	1.25E-11	6.41E-12	2.45E-11
0.285	8.23E-07	3.73E-11	2.00E-11	6.96E-11
0.290	5.20E-07	1.08E-10	5.82E-11	1.94E-10
0.295	3.27E-07	3.07E-10	1.74E-10	5.49E-10
0.300	2.04E-07	8.47E-10	4.84E-10	1.46E-09
0.305	1.27E-07	2.28E-09	1.35E-09	3.85E-09
0.310	7.86E-08	5.97E-09	3.69E-09	9.81E-09

Standardization Issue

Since there are many ways of implementation for finger vein biometric algorithm as described above, it is very important to standardize the basic framework of the biometric system in order to expand and guarantee the compatibility. There are many ongoing projects and activities aiming to standardize various biometric modalities. One of the most comprehensive and widely-recognized groups is the Sub Committee 37 (SC37) of the Joint Technical Committee for Information Technology (JTC1) [5]. JTC1 is a joint project established by the International Organization for Standardization (ISO) [6] and the International Engineering Consortium (IEC) [7]. SC37 is dedicated to the standardization of biometrics since 2002 and is one of the 18 active Sub Committees of the joint project. SC37 members are all national bodies, and there are 25 participating countries and 7 observing countries as of October 2007. SC37 has already released 20 official standards including a standard for biometric vascular image data published in 2007 [8].

Summary

Although finger vein biometrics is one of the latest biometric modalities, its feature extraction algorithm has been continuously improved since the beginning of its fundamental research in early 1990s. The feature extraction algorithm described in this document is based on one of a very few academic papers reporting the core part of the finger vein biometrics; however, it is quite possible that the feature extraction methods employed by commercially available products today have already been modified or totally renewed. This continuous improvements and updates of the algorithm are, in many cases, beneficial or even preferable from a security point of view. Finger vein biometrics is with no doubt one of the most accurate biometric modalities available today. With its high usability and user-acceptability, it is highly anticipated that this new biometric technology will establish a de facto standard of the next generation access control system in various application fields.

Related Entries

- ▶ Finger Vein
- ▶ Finger Vein Imaging
- ▶ Finger Vein Reader

References

1. Hitachi, Ltd. http://www.hitachi.com/
2. Miura, N., Nagasaka, A., Miyatake, T.: Feature Extraction of Finger Vein Patterns Based on Iterative line Tracking and Its Application to Personal Identification. IEICE Trans. Inf.Syst. **J86-D-II**(5), 678–687 (2003) (Japanese Edition)
3. Kono, M., Ueki, H., UmemuraS.: Near-infrared finger vein patterns for personal identification. Appl. Opt. **41**(35), 7429–7436 (2002)
4. Yanagawa, T., Aoki, S., Ohyama, T.: Human finger vein images are diverse and its patterns are useful for personal identification MHF Preprint Series, MHF 2007-12, Kyushu University 21st Century COE Program, Development of Dynamic Mathematics with High Functionality (2007)
5. The Joint Technical Committee for Information Technology: www.jtc1.org
6. The International Organization for Standardization:http://www.iso.org/iso/home.htm
7. The International Engineering Consortium:http://www.iec.org/
8. ISO/IEC 19794-9:2007:Information technology – Biometric data interchange formats – Part 9: Vascular image data. (2007)

Finger Vein Feature Segmentation

► Finger Vein Biometric Algorithm

Finger Vein Imaging Device

► Finger Vein Reader

Finger Vein Pattern Imaging

Mitsutoshi Himaga
Hitachi-Omron Terminal Solutions, Corp., Tokyo, Japan

Definition

A technology to visualize and capture an individual's finger vein network patterns.

Introduction

Blood vessels are not exposed out of the human body and its network patterns are normally impossible to see without the range of visible light wavelength (Retinal blood vessels are the only exception, which can be seen in visible light. However, it is necessary to use specially designed devices such as ophthalmoscopy or retinal scanner to observe the blood vessels on retina.). In order to visualize blood vessel patterns that are hidden under the skin, it is necessary to use appropriate imaging technologies. It is well known that hemoglobin absorbs ► *near infrared rays* more than other substances that comprise human body. Since most of the hemoglobin in human body exists in red blood cells that are flowing inside blood vessels, the blood vessel network patterns can be seen as dark area by infrared imaging systems. Finger vein pattern imaging is a technology that utilizes this optical characteristic of hemoglobin, by which vascular network patterns inside the finger of an individual are visualized. The raw images taken by using infrared lights can further be improved by appropriate illumination control and image processing techniques such as contrast enhancement so that biometric information can be extracted. Although the same sort of technology is widely used in medical fields (which are sometimes referred to as optical coherence tomography or OCT), the scope of this document is limited to its biometric applications only.

Infrared lights projected in a human body can easily be diffused and the contrast of blood vessels and the background is rapidly deteriorated as the infrared light penetrates deeper into the part of the body. This is sometimes compared to a swizzle stick put in a glass of milk. The swizzle stick can be seen from outside when it is close to the interior surface of the glass, however, it becomes gradually invisible when it is moved towards the middle of the glass due to the light diffusion. Therefore, it is believed that the vascular network patterns visualized by infrared illumination exist in the area that is close to the skin. Considering the resolution of the cameras commonly used for finger vein biometrics and the fact that the diameters of arteries are as small as approximately 1/3 of those of veins in finger, it is reasonable to assume that most of the visualized blood vessels are veins. This is why many of vascular biometric technologies are known as "vein" biometrics, though arteries and veins are equally visualized by infrared light and normally treated in the same manner.

Light Source

The most commonly used light source for blood vessel pattern imaging is infrared light emitting diodes (IR-LEDs). The IR-LED is not a newly developed product; they are being widely used for household appliances such as TV remote controllers for a long time, which proves the safety for human beings and livestock. In the actual implementation, there are many forms of the light source arrangements depending on the target. Finger vein imaging systems typically require small and oblong field of view, and therefore linear arrays of IR-LEDs are usually preferred. On the other hand, grid or circular light source arrangements are more appropriate for the systems that require larger field of view. Many of palm vein and back-of-hand biometric systems employ this type of light source configuration.

Illumination Control

Finger vein patterns are distinct from other biometric features as they are inside human body and unnoticeable. This is, of course, one of the major advantages of the biometric modality, however, it is also a big challenge to capture a clear finger vein image. Since the finger vein network has a three-dimensional structure, some parts are close to the skin surface and others are not. This makes it very difficult to obtain high and homogeneous image contrast throughout the region of interest. Furthermore, the thickness of finger has a large individual variation, which results in a variety of distances between the finger and the LED arrays. Therefore, it is almost obvious that there is no single perfect illumination setting that accommodates all these variations and this is why the illumination control technology is considered to be one of the key factors of the finger vein biometrics.

At the time of authentication process, it is virtually impossible to obtain an image that is pixel-wise identical to the enrolled pattern due to the differences caused by the change in environment or the positioning of the sample. If only one sample image is to be matched to the template per attempt, it is likely to have very high false rejection rate (FRR). In order to cope with this difficulty, most of vein biometric systems continuously capture the presented sample with several illumination configurations. Each of the captured vein patterns is matched to the template one by one in real time, and the system continues this loop until the presented sample is either accepted or rejected. Therefore, it is very important to design the illumination control algorithm to produce optimized images as quickly as possible so that genuine attempt can be processed in a short time. The details of the illumination control algorithms are, however, confidential in most cases, and not published by any vendors at the time of writing.

The Imaging Methods

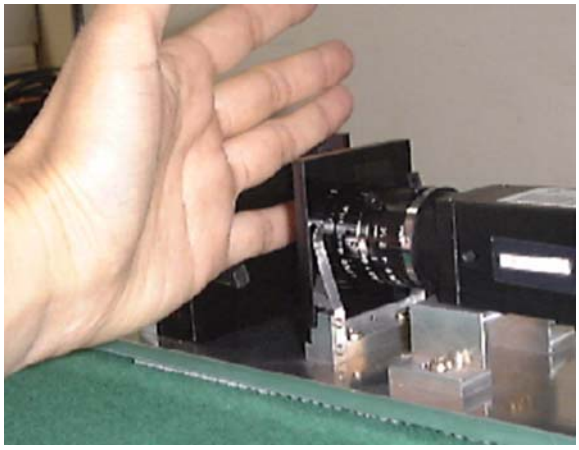
There are two major approaches to visualize vascular patterns for biometric use, the light penetration method and the light reflection method. The light penetration method utilizes the infrared light transmitted through the target object, while the light reflection method makes use of the light reflected by the target. The light reflection method is not usually a first choice unless it is necessary (e.g., retinal blood vessel patterns) because it is difficult to handle the reflected images that may contain saturated (over-exposed) areas or texture on the skin surface. The contrast of the images captured by penetrating light is generally higher than that by reflected light; and therefore, most commercially available finger vein biometric systems employ the light penetration method.

We will focus on the finger vein imaging technologies based on the light penetration method in this essay.

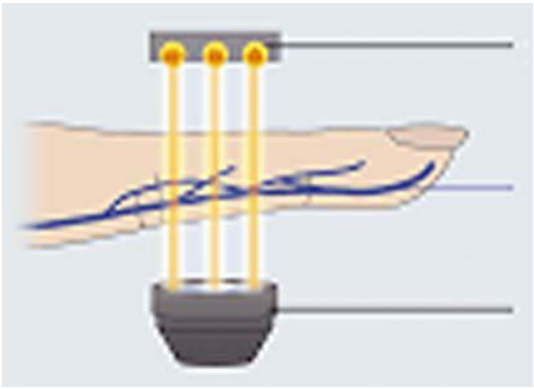
There are three major implementations of the finger vein imaging system. In the following part of this essay, the three finger vein imaging systems are briefly reviewed in chronological order along with some examples of its commercial products and applications.

Top-lighting Systems

Finger vein readers whose infrared light source is placed on the other side of the camera with respect to the finger are called top-lighting systems. Hitachi Central Research Laboratory (Tokyo, Japan) [1] started the research and development of finger vein biometrics in mid-1990's [2] and evaluated the technology by using a prototype of this lighting system (Fig. 1). As illustrated in Fig. 2, infrared rays are projected from the opposite side of the infrared camera with respect to the sample finger, which visualize the finger vein patterns on the camera side.



Finger Vein Pattern Imaging. **Figure 1** Finger vein imaging systems(prototype) (Courtesy of Hitachi, Ltd.).



Finger Vein Pattern Imaging. **Figure 2** Top-lighting system (Courtesy of Hitachi, Ltd.).

The top-lighting imaging method has the following features.

- Robust against environmental illumination.
  The light source housing protects the camera from unwanted ambient lights that deteriorate the image quality. This structure makes the top-lighting system the most robust imaging system in terms of environmental changes.
- Stable illumination.
  Since the top-lighting system has only one light source placed right behind the finger, the contrast attenuation of the captured image is isotropic and no special image processing is required as long as the region of interest has enough signal-to-noise ratio.

Since this is the earliest and the most straightforward implementation of finger vein imaging system, many commercial models today employ this approach for both logical and physical access control applications. One of the earliest commercial finger vein products is a physical access control system developed by Hitachi Engineering Co., Ltd. (Its biometrics division was reorganized into Hitachi Information and Control Solutions, Ltd. in 2006 [3].) in 2002. Their product, SecuaVeinAttestor® employed the top-lighting system and demonstrated very stable performance. This product was further improved in terms of robustness in the following year and achieved even higher accuracy comparable to iris recognition (Fig. 3). Figure 4 shows a logical access control unit PC-KCA100 jointly developed by Hitachi, Ltd. (Tokyo, Japan) [4] and Hitachi Software Engineering, Co., Ltd. (Tokyo, Japan) [5] in 2006. This product has an application programming

Finger Vein Pattern Imaging. **Figure 3** SecuaVeinAttestor® (Courtesy of Hitachi Information & Control Solutions, Ltd.).

Finger Vein Pattern Imaging. **Figure 4** Hitachi PC-KCA100 (Courtesy of Hitachi, Ltd.).

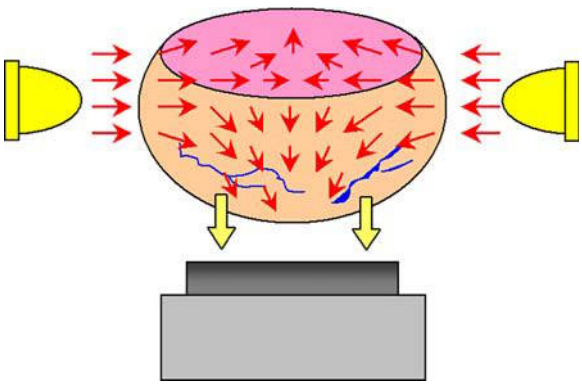
interface (API) that was developed based upon ► *BioAPI*., which enables the biometric device to easily communicate with many types of systems. Another interesting example of top-lighting system was demonstrated by Hitachi, Ltd. in 2007. It introduced a unique automobile ignition key device (prototype) in Tokyo Motor show 2007, which allows pre-enrolled drivers to start the engine by presenting their fingers on the finger vein reader embedded on the steering wheel [6].

Side-lighting Systems

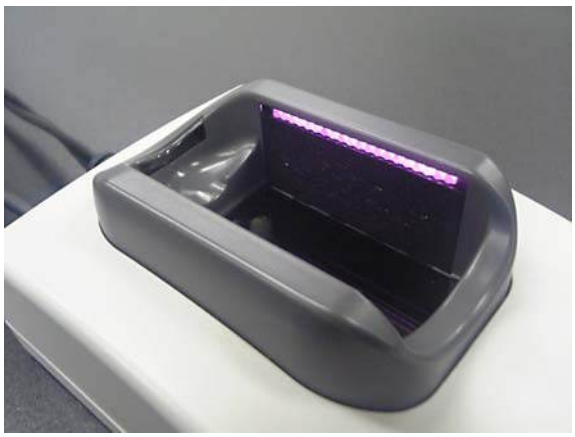
Side-lighting systems typically have a pair of infrared LED arrays embedded on the both sides of the presented finger. The infrared rays emitted by the light source propagate inside the finger and some of them reach to the infrared camera placed beneath the finger as illustrated in Fig. 5. Figure 6 shows an example of the side-lighting system.

This imaging method has the following features.

- Medium-sized enclosure
- User-friendly design; low psychological barrier



Finger Vein Pattern Imaging. **Figure 5** Side-lighting system (Courtesy of Hitachi-Omron Terminal Solutions, Corp.).



Finger Vein Pattern Imaging. **Figure 6** Infrared LED array (Courtesy of Hitachi-Omron Terminal Solutions, Corp.) Infrared LEDs are colored in this picture for visualization.

Unlike the top-lighting systems, the presented finger is always within the field of view of the user, which considerably reduces psychological difficulties of the user while scanning.

- High maintainability

It is easy to clean up the camera surface because no housing covers the optical unit.

Although the side-lighting systems require very advanced image processing and illumination control technologies, it is one of the most popular implementations that is employed by many commercial models. One of the most widely used applications of this lighting system is automated teller machines (ATMs). Hitachi-Omron Terminal Solutions, Corp. (Tokyo, Japan) [7] is the only supplier of finger vein authentication systems for banking transactions as of 2007, who has shipped approximately 40,000 ATMs equipped with finger vein biometrics (Fig. 7) and enrollment units (Fig. 8) in Japan since 2005. Hitachi-Omron has also developed a unique key management system with finger vein authentication in 2006 (Fig. 9). Hitachi Software Engineering, Co., Ltd. developed a compact logical access control unit called Johmon J200 in 2004, which employs the side-lighting system.

Bottom-lighting Systems

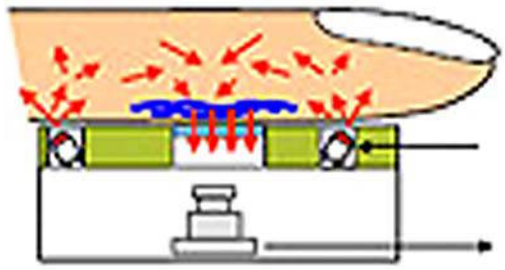
Bottom-lighting systems have been developed as an answer to the growing demand for mobile applications. Typically, the bottom-lighting systems have a

Finger Vein Pattern Imaging. **Figure 7** ATM equipped with a finger vein reader (Courtesy of Hitachi-Omron Terminal Solutions, Corp.).

Finger Vein Pattern Imaging. **Figure 8** Hitachi-Omron's UBReader (Courtesy of Hitachi-Omron Terminal Solutions, Corp.).

Finger Vein Pattern Imaging. **Figure 9** Key management system with a finger vein reader (Courtesy of Hitachi-Omron Terminal Solutions, Corp.) Metal enclosure is removed for demonstration purpose.

pair of infrared LED arrays and an infrared camera embedded on the same surface as shown in Fig. 10. Although the configuration of the components is similar to the imaging systems using reflection light such as palm vein readers, the finger has to be touched to the LED arrays while scanning. The infrared rays projected into both the tip and the root of the presented finger propagate inside the finger and visualize



Finger Vein Pattern Imaging. **Figure 10** Bottom-lighting system (Courtesy of Hitachi, Ltd.).

the vascular patterns in the same manner as the side-lighting systems. This imaging method has the following features.

- Cost effective
- Very small in volume

  Since both the light source and the camera are embedded on the same surface, it does not require any three-dimensional structure. This enables the imaging system flexibly embedded to many devices including portable devices such as laptop computers or mobile phones. The volume of Hitachi's prototype unit developed in 2005 [8] is as small as 19 ml (39 mm (depth) × 34 mm (width) × 15 mm (height)), and further miniaturization is expected in the near future.
- User-friendly design; minimum psychological barrier

  Since the bottom-lighting systems can be embedded to existing hardware without changing the original appearance of the hardware significantly, user's acceptability is the best among the three systems described here. The exterior of the scanning unit is quite similar to the widely used finger print scanners and thus psychological barrier of the user is very low.
- High maintainability

  Cleaning the bottom-lighting imaging system is as easy as wiping a flat surface. In addition, it is not necessary to clean the system as frequent as other two systems because it has no holes or ditches in which dust can accumulate.

Hitachi, Ltd. released the first laptop PC equipped with an embedded finger vein authentication module in 2005 by using this imaging technology (Fig. 11).

Finger Vein Pattern Imaging. Figure 11
Bottom-lighting system (Hitachi Laptop PC Se210) (Courtesy of Hitachi, Ltd.).

Summary

As described in this essay, finger vein imaging systems available today can be categorized into three groups: the top-lighting systems, the bottom-lighting systems and the side-lighting systems. Because each system has its unique features, it is very important to choose a suitable imaging system depending on the application. In general, the reproducibility of imaging systems is, to some extent, subject to environmental changes such as ambient lightings or the conditions of subject, and of course, none of the three imaging systems overviewed here are free from these constraints. In other words, these changes can be regarded as external (uncontrollable) parameters and imaging systems that are robust against these parameters are generally preferred. The performance of a biometric system can be improved by suppressing the influence of these uncontrollable parameters as little as possible; and it is particularly important to select an appropriate imaging system depending on the application by taking the variety and the range of these parameters into consideration.

Related Entries

- ▶ Finger Vein
- ▶ Finger Vein Feature Extraction
- ▶ Finger Vein Reader

References

1. Hitachi Central Research Laboratory, http://www.hqrd.hitachi.co.jp/crle/
2. Kono, M., Ueki, H., Umemura, S.: Near-infrared finger vein patterns for personal identification. Appl.Opt. **41**(35), 7429–7436 (2002)
3. Hitachi Information and Control Solutions, Ltd., http://www.hitachi-ics.co.jp/product/english/index_en.htm
4. Hitachi, Ltd., http://www.hitachi.com/
5. Hitachi Software Engineering, Co., Ltd., http://hitachisoft.jp/English/index.html
6. Hitachi develops finger vein authentication technology for steering wheels. Hitachi, Ltd. News release, Oct. 2007. http://www.hitachi.com/New/cnews/071022b.html
7. Hitachi-Omron Terminal Solutions, Corp., http://www.hitachi-omron-ts.com/index.html
8. Hitachi develops compact finger vein authentication technology for laptop PCs. Hitachi, Ltd. News release, Oct. 2005. (in Japanese) http://www.hitachi.co.jp/New/cnews/month/2005/10/1003.pdf

Finger Vein Reader

Mitsutoshi Himaga
Hitachi-Omron Terminal Solutions, Corp. Tokyo, Japan

Synonyms

Finger vein scanner; Finger vein imaging device; Finger vein authentication device

Definition

A finger vein reader is a biometric device that comprises at least one optical imaging unit designed to capture finger vein patterns of an individual and a digital signal processor that digitizes the captured finger vein patterns to be utilized as biometric features.

Introduction

Unlike conventional biometric features such as finger print, vascular network patterns cannot be observed without using specially designed equipment and thus, it is extremely difficult to steal or duplicate the biometric information. The possibility of biometric

identification based on human finger vein patterns captured by transmitting light was indicated by Shimizu in 1992 [1]. However, it was not until Kono et al. developed a near-infrared finger vein reader prototype and demonstrated its effectiveness in 2000 that the concept became reality [2]. Kono et al. further evaluated the performance of the proposed biometric modality by using sample data collected from 678 subjects and reported very positive results in 2002 [3]. In 2007, Yanagawa et al. demonstrated the diversity of human finger vein patterns by conducting statistical analysis based on sample data collected from 506 subjects. They also proved the feasibility of using finger vein patterns for personal identification by evaluating false acceptance rates (FAR) and false rejection rates (FRR) [4]. Today, finger vein biometrics is recognized as one of the most reliable and secure biometric modalities and is applied to a variety of security systems.

Features of Finger Vein Modality

The advantages of finger vein biometrics are summarized as below:

1. Accuracy

   Finger vein biometrics is one of the most accurate biometric modalities available today. A finger vein authentication device called UBReader has been certified as level 3 in the accuracy scale by the US-based International Biometric Group [5, 6]. No other biometric device has been rated at the highest possible level, level 4. The details of the evaluation results are reported in the Comparative Biometric Testing (CBT) round 6 Public Report [7].
2. Usability

   Finger vein biometrics can be implemented in many forms according to the demands and requirements of the application. This flexibility makes it possible to design the hardware optimized to a specific use. For example, Hitachi-Omron's UBReader, which was primarily designed for banking application, demonstrated very high usability in terms of indices such as ▶ Failure-to-Enroll Rate (FTE) or ▶ Enrollment Transaction Duration in the CBT and achieved level 3 in the usability scale of the testing.
3. Compactness/Flexibility

   Since the target imaging area of a finger vein reader is generally smaller than for other vascular pattern biometric devices (e.g., palm vein or the back-of-the-hand vein systems), finger vein readers can be installed into a variety of devices flexibly. One of the most compact finger vein readers was not more than 19ml in volume, which made possible for laptop PCs to embed the device without changing their appearances. The short focus depth (i.e., the distance between the camera and the target) makes it easy to align the finger and, therefore, no hand-guide or handle bars, which are sometimes necessary for other hand vascular devices, are needed.
4. Small templates

   The size of finger vein template is typically some hundreds of bytes per finger. This means that finger vein biometric database can be very cost-effective because it does not require a large storage system, compared with other biometrics. This feature is also preferable for systems which store templates on a server and transmit them upon request over a network. Small template size makes a big difference especially when a high-speed network is not available or the data traffic is very high.
5. Excellent image quality

   Since the raw image is the very first input from which most biometric information is extracted, the image quality is largely responsible for the overall performance of the biometric system. All finger vein readers, commercially available today, utilize near infrared rays that are projected through the presented finger. The images captured by using this method (known as the "light penetration method") have very high contrast and little noise because most of ridges and wrinkles on the skin are not imaged.
6. More back-up samples

   Unlike most biometric systems, finger vein biometrics allows more than two templates per person. Even in the case when one of the enrolled fingers gets injured and cannot be presented to the biometric system, it is possible to operate the system using other fingers.

The Hardware

Finger vein readers can be classified into three different groups, depending on the device used, and execution of the enrollment and authentication processes.

The finger vein readers in the first category do not have an authentication algorithm on the readers which

is known as "match-on-PC" readers. Instead, the authentication algorithm is implemented as a computer software and distributed together with the finger vein reader. The software is installed to the host PC beforehand, where the enrollment and authentication processes are executed. Since the match-on-PC finger vein readers do not need a powerful processor, the cost of the hardware is relatively low compared with the other two kinds of finger vein readers. Due to the low power consumption, most of the match-on-PC devices can be driven by the 5 volts power supplied through the universal serial bus (USB) interface, which contributes to the compactness and the portability of the device. Since the turn-around time of the authentication process is dependent on the host PC's CPU power and the communication speed of the interface, the throughput of the entire system may vary. Although the match-on-PC readers are widely used for the purpose of logical access control (e.g., PC log-in), they are increasingly coming into use for physical access control applications.

The second category is called "match-on-device" finger vein readers. The match-on-device reader is equipped with a CPU that executes both enrollment and authentication processes inside the reader itself. The authentication algorithm is implemented in firmware and is typically encrypted when stored on a non-volatile static memory. One of the biggest advantages of this system is that all algorithms and data required for biometric authentication are enclosed in a ▶ tamper-proof casing and completely separated from the outside world. Since all biometric data and algorithms can be stored inside of the finger vein reader, the risk of hacking is minimal. Another advantage of this system is that the match-on-device finger vein readers do not require high-performance host PCs. In most cases, a low-performance CPU is enough to communicate and control the match-on-device finger vein reader, which makes it possible to integrate cost-effective systems. The data communications between the host PC and the finger vein reader are limited because no biometric data is needed to transfer and therefore no high-speed interface/network is required. The unit price of these readers tends to be higher than the match-on-PC readers; however, the match-on-device readers can be used for a wide range of applications as they are suitable for both high-security systems and low-cost systems. Typical applications of the match-on-device readers include banking systems and physical access control systems.

The third category is known as ▶ match-on-card finger vein readers. The authentication algorithm is implemented as smart card application software and securely stored onto a smart card together with biometric templates. Upon the host PC's request, the match-on-card finger vein reader extracts the biometric feature of the presented finger and sends an authentication command to the smart card together with the features. The smart card then executes the authentication algorithm on its own CPU embedded inside and evaluates the features transmitted by the finger vein reader. After the smart card determines whether the presented finger matches with the pre-enrolled template, it transmits a response back to the host PC through the reader. One of the benefits of using the match-on-card system is its high security feature. Both the authentication algorithm and the template data are securely stored on a smart card that is inaccessible without taking validation procedures using Secure Application Module (SAM). Since these data is never transmitted outside the card, the risk of template duplication is extremely low. From a viewpoint of system administration, the risk management cost of the match-on-card system can be dramatically suppressed because the system does not need to provide protection for the template data (the card holders are responsible for their own templates, instead). Though the authentication processing time is slightly longer than other two kinds of readers (this is because the smart card CPUs are slower than the embedded CPUs or PCs), it does not make much difference especially for its primary usage, verification. For these reasons, match-on-card finger vein systems are currently the most popular biometric banking solution in Japan.

Security Features

Some finger vein readers have a security measure called ▶ liveness detection. It is very important for biometric systems in general to ensure that the enrolled biometric patterns are genuine. If a biometric device accepts any artifact mistakenly and enroll it as a genuine template, that can be used just like a normal key that can be used by anyone; if this happens, the security level of the biometric system becomes no higher than conventional keys and locks. In the actual applications, enrollment procedures typically require an administrator to be present (who will never allow users to enroll artifacts);

however, it is still beneficial to have this security measure because it is also used in the authentication procedure in order to ensure that the presented sample is from a live body. Liveness detection can be implemented by either hardware or software (or both) and there are many different methods to realize the functionality. The details of the method employed by finger vein readers are, however, not publicly available at the time of writing due to the secure nature of the functionality.

Another security feature that some finger vein readers have is the tamper-proof structure. This structure enables the system to identify that it has been tampered with, and in some cases, to disable itself when unauthorized person try to dismantle or reverse-engineer the system. This security measure is especially important when the biometric system is to be used by open public, for instance, ATMs. Just like liveness detection, the details of the tamper-proof structure are highly confidential and no finger vein manufacturer discloses the mechanism for security reasons.

Finger Vein Reader. **Figure 1** Finger vein reader implemented on an ATM.

Applications

- Banking transactions

  Banking applications are currently the most popular application of finger vein biometrics. The first finger vein biometric ATM system was developed and introduced by Hitachi-Omron Terminal Solutions, Corp. in 2005. The biometric ATM was equipped with an open-scanning finger vein reader, as shown in Fig. 1, and adopted by one of the largest banks in Japan, Sumitomo Mitsui Banking Corporation (SMBC, Tokyo, Japan) [8] and later, was widely adopted by more than 60 financial institutions in Japan including Japan Post Bank Co., Ltd. [9, 10, 11]. According to a recent survey more than 80% of Japanese financial institutions that adopted biometric banking systems employ finger vein biometrics [12]. It is expected that more than 40,000 ATMs in Japan will be equipped with finger vein readers by the end of 2008, which will make up approximately 25% of ATMs of the country.

  In typical finger vein banking systems, each account holder who wishes to have his or her biometric data enrolled visits a branch of the bank in person and enrolls two fingers at the teller counter after prescribed personal identification procedure. The templates are then stored in a smart card issued by the bank, on which the matching process is executed during the authentication process (i.e., "*match-on-card*" technology). Since the matching process is executed against the two templates stored on the smart card, users can present either of the two enrolled fingers. Many of the finger vein ATM networks are connected to each other and the account holders can use their biometric bankcards at any ATM that belongs to the participating financial institutions.

- Door access control

  Door access control is another popular application of the finger vein biometrics. The first commercial application of the finger vein biometrics was a door access control system called SecuaVeinAttestor® developed by Hitachi Engineering Co., Ltd. in 2002. (Please note that Hitachi Engineering Co., Ltd. reorganized its biometrics division into Hitachi Information and Control Solutions [13], Ltd. in 2006.) The door access control system is equipped with a ten-key pad, with which users type his or her ID number so that it can execute one-to-one matching (verification). It can also be used with proximity cards, which allow users to unlock the door without typing their ID numbers. In addition, a biometric door access control system has been developed that works with electric locks [14, 15]. A prototype automobile entry system using finger vein biometrics was demonstrated in the Tokyo Motor Show in 2005, which enables pre-registered users unlock the door just by holding the door handle (Fig. 2).

Finger Vein Reader. Figure 2 Finger vein reader embedded on a door handle.

- Logical access control

  Logical access control is also a popular and widely used application. Since host computers (PCs) are normally equipped with a CPU powerful enough to execute the matching process in real time, many finger vein readers for this application employ the match-on-PC architecture. Hitachi, Ltd. and Hitachi Software Engineering (Tokyo, Japan) [16] jointly developed a very compact finger vein reader for PC called PC-KCA100 in 2006. This match-on-PC finger vein reader has an application programming interface (API) based on the widely recognized international standard BioAPI 2.0, which enables it to easily communicate with many types of systems. The power consumption of PC-KCA100 is so small (less than 2.5 watts) that it can be driven by the power supplied by the USB interface only.
- Other applications

  Amano Corporation (Kanagawa, Japan) [17] developed the first "time and attendance" terminal equipped with a finger vein reader called AGX250AV in 2007. This innovative terminal can store up to 1000 finger vein templates and authenticate the users without using an ID card. In addition to the convenience, AGX250AV eliminates inappropriate attendance records by impostors (this is known as "buddy punching"), which dramatically increases the reliability of the time information system. Alpha Locker System Co., Ltd. (Kanagawa, Japan) [18] developed the first finger vein biometric locker FB-BM in 2007. The biometric locker, which is aimed for public use, has some tens of doors that can be accessible by presenting a finger. FB-BM is capable of identifying a pre-enrolled finger by using one-to-many matching algorithm and does not require the users to specify which door to open before presenting their fingers.

Summary

Although finger vein biometrics is one of the latest biometric technologies, it has already established both technical and statistical feasibility. Finger vein readers have been successfully applied to a growing array of applications such as time and attendance or physical access control systems. Its banking applications remain one of the largest and the most successful set of applications for this state-of-the-art biometric modality; and it is anticipated that more than a quarter of ATMs in Japan will be equipped with finger vein readers by the end of 2008. Some financial institutions who adopted other hand vascular biometrics started to modify their systems to accept finger vein biometric data, or even replace their systems with finger vein readers. This trend is expected to continue as the number of finger vein readers increase, and it is very likely for the biometric technology to set a new standard in security applications in the very near future.

Related Entries

- ▶ Finger Vein
- ▶ Finger Vein Feature Extraction
- ▶ Finger Vein Imaging

References

1. Shimizu, K.: Optical trans-body imaging: feasibility of optical CT and functional imaging of living body. Jpn. J. Medicina philosophica **11**, 620–629 (1992) (in Japanese)
2. Kono, M., Ueki, H., Umemura, S.: A new method for the identification of individuals by using vein pattern matching of a finger. In: Proceedings of the Fifth Symposium on Pattern Measurement, pp. 9–12, Yamaguchi, Japan. (2000) (in Japanese)
3. Kono, M., Ueki, H., Umemura, S.: Near-infrared finger vein patterns for personal identification. Appl. Opt. **41**(35), 7429–7436 (2002)

4. Yanagawa, T., Aoki, S., Ohyama, T.: Human finger vein images are diverse and its patterns are useful for personal identification. MHF Preprint Series, MHF 2007–12, Kyushu University 21st Century COE Program, Development of Dynamic Mathematics with High Functionality (2007). http://www2.math.kyushu-u.ac.jp/coe/report/pdf/2007-12.pdf
5. UB Reader is developed by Hitachi-Omron Terminal Solutions, Corp. http://www.hitachi-omron-ts.com
6. International Biometric Group, LLC. http://www.biometricgroup.com
7. Theme, M., ed.: Comparative Biometric Testing Round 6 Public Report. The International Biometric Group (2006)
8. Sumitomo Mitsui Banking Corporation, http://www.smbc.co.jp/global/index.html
9. Japan Post Bank Co., Ltd. http://www.jp-bank.japanpost.jp/en_index.html
10. Mizuho Bank, Ltd. http://www.mizuhobank.co.jp/english/
11. Resona Bank, Limited. http://www.resona-gr.co.jp/holdings/english/index.html
12. Conducted by Hitachi-Omron, January 2008
13. Hitachi Information and Control Solutions, Ltd., http://www.hitachi-ics.co.jp/product/english/index_en.htm
14. Electric locks developed by Miwa Lock Co., Ltd. http://www.miwalock.com/
15. Hitachi, Ltd. http://www.hitachi.com/
16. Hitachi Software Engineering, Co., Ltd.http://hitachisoft.jp/English/index.html
17. Amano Corporation, http://www.amano.co.jp/English/index.html
18. Alpha Locker System Co., Ltd., http://www.alpha-locker.com/index.html

Finger Vein Scanner

▶ Finger Vein Reader

Fingermark Identification Procedure

▶ Fingerprint, Forensic Evidence of

Fingerprint

Fingerprint is an impression or image left on a surface by the friction skin of a finger.

▶ Anatomy of Friction Ridge Skin

Fingerprint Analysis

▶ Fingerprint Features

Fingerprint Authentication

▶ Fingerprint Indexing

Fingerprint Benchmark

▶ Fingerprint Databases and Evaluation

Fingerprint Binarization

Fingerprint binarization is the process of converting an 8-bit gray-scale fingerprint image into a 1-bit ridge image. This is virtually equivalent to thresholding. Post-processing for the binarized image, such as smoothing, is also important.

▶ Fingerprint Image Enhancement

Fingerprint Biometric

▶ Fingerprint Recognition, Overview

Fingerprint Capture

▶ Biometric Sample Acquisition

Fingerprint Characteristics

► Fingerprint Features

Fingerprint Classification

Xudong Jiang
Nanyang Technological University, Nanyang Link, Singapore

Synonyms

Fingerprint indexing; Fingerprint pre-matching; Fingerprint retrieval

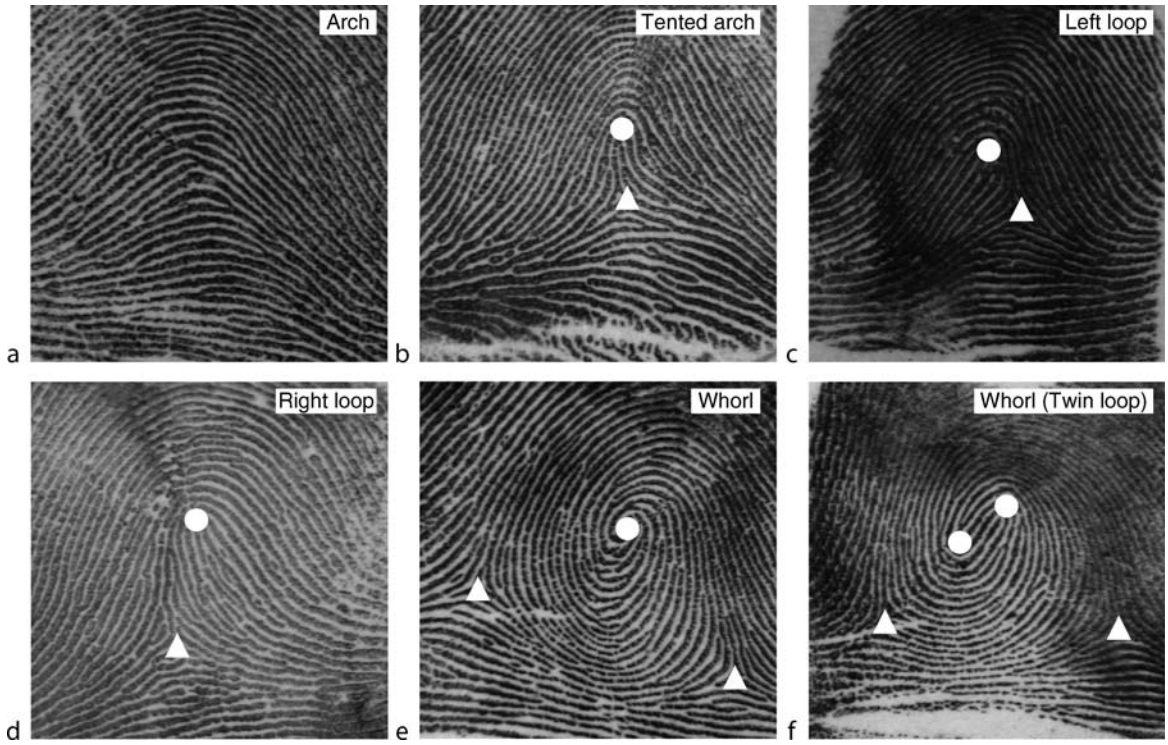
Definition

Fingerprint classification is a procedure in which fingerprints are grouped in a consistent and reliable way, such that different impressions of a same finger fall into a same group. It can be viewed as a coarse-level pre-matching procedure so that a query fingerprint needs to be further compared with only a smaller subset of fingerprints in the database belonging to the same group. It is often necessary to integrate a classification module into a fingerprint identification system to speed up the database search. A database can be partitioned into ► human-interpretable fingerprint classes based on Galton–Henry scheme or into ► machine-generated fingerprint classes.

Introduction

A fingerprint recognition system captures a user's fingerprint and compares it with the information stored in a database to establish or to authenticate his/her identity. If an identity is claimed, the system compares the query fingerprint only with the template corresponding to this identity stored in the database. This one-to-one matching process is called fingerprint verification. If no identity is claimed, the system needs to compare the query fingerprint with all templates stored in the database to establish the identity. This one-to-many matching process is called fingerprint identification. The extension of the one-to-one matching of a verification system to the one-to-many matching of an identification system increases the possibility of false positive matching. Comparing to the verification performance, both accuracy and speed may deteriorate significantly if a verification algorithm is naively extended to solve an identification problem. The performance deterioration could be very serious for large-scale identification systems as it is directly proportional to the number of fingerprints in the database [1]. This problem can be alleviated by reducing the search space of exact matching. Fingerprint classification, indexing, or retrieval techniques facilitate the reduction of the search space. They can be viewed as a coarse-level pre-matching process before further exact matching in an identification system. A query fingerprint is first compared to prototypes of the pre-specified classes, bins or clusters to find its class membership. Then, it is only necessary to compare the query fingerprint exactly with a subset of the database that has the same class membership. For example, if a database is partitioned into ten groups, and a query fingerprint is matched to two of the ten prototypes, then the identification system only needs to search two of the ten groups of the database for exact matching. This reduces the search space by fivefold if fingerprints are uniformly distributed in the ten groups.

The first rigorous scientific study on fingerprint classification was made by Sir Francis Galton in the late 1880s [2]. Classification was introduced as a means of indexing fingerprints to speed up the search in a database. Ten years late, Edward Henry refined Galton's work and introduced the concept of fingerprint "core" and "delta" points for fingerprint classification [3]. Figure 1 shows the five most common classes of the Galton–Henry classification scheme where the core and delta points and the class names are shown. Henry's classification scheme constitutes the basis for most modern classification schemes. Most law enforcement agencies worldwide currently employ some variants of this Galton–Henry classification scheme. Although Galton–Henry scheme has some advantages, such as human-interpretable and rigid segmentation of a database, only a limited number of classes are applicable to the automated system. For example, most automated systems [4–8] can only classify fingerprints into five classes as shown in Fig. 1. Moreover, fingerprints



Fingerprint Classification. Figure 1 Six sample fingerprints from the five commonly used fingerprint classes (arch, tented arch, left loop, right loop, and whorl) under the Galton–Henry classification scheme where two whorl fingerprints are shown (a plain whorl and a twin loop whorl). Singular points of the fingerprints, called core and delta, are marked as *filled circles* and *triangles*, respectively. Note that fingerprints of an arch class have neither core nor delta.

are not evenly distributed in these classes and there are some ambiguous fingerprints that cannot be reliably classified even by human experts. Therefore, Galton–Henry scheme that partitions the database into human-interpretable fingerprint classes is not immune to errors and does not offer much selectivity for fingerprint searching in large databases.

In fact, it is not obligatory for an automated system to partition the database into human-interpretable fingerprint classes. In automatic fingerprint identification systems (AFIS), the objective of the classification is to reduce the search space. This objective can be accomplished by partitioning the database into machine-generated fingerprint classes in feature space as long as the classification is consistent and reliable. For example, some fingerprint index techniques [9, 10] can reduce the search space more efficiently than the Galton–Henry scheme. Continuous classification techniques [1, 11, 12] do not pre-classify the database, but represent each fingerprint with a numerical feature vector. Given a query fingerprint, a class is formed by retrieving a portion of fingerprints from database whose feature vectors are close to that of the query fingerprint. Although these techniques can classify fingerprints into large number of classes, a query fingerprint needs to be compared with all fingerprints in the database, which could be time consuming for a large database. This problem can be circumvented by incorporating data clustering techniques in the ▶ fingerprint retrieval framework [12, 13].

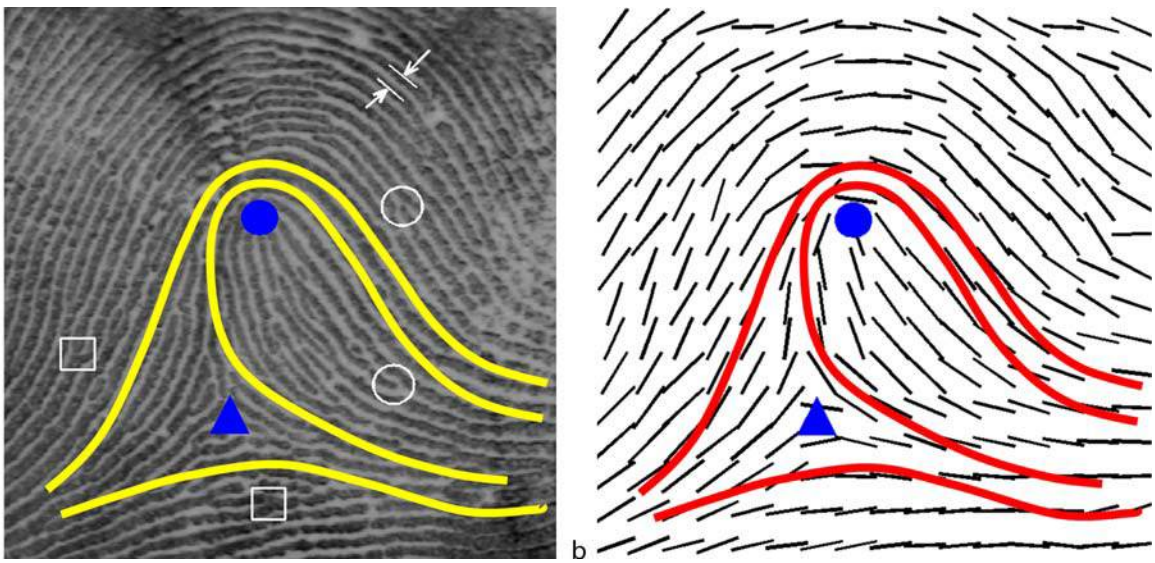
Feature Extraction for Classification

Not all measurements of a fingerprint image remain invariant for a given individual over the time of capture and can be used to discriminate between identities. The first step of fingerprint classification is to find

salient features that have low intra-class variation and high inter-class variation. Fingerprint image is an oriented texture pattern that contains ridges separated by valleys and exhibits two levels of feature as shown in Fig. 2. At the global level, the orientation field and the ridge frequency are two primitive and fundamental features. At the fine local level, the most prominent characteristics are the minutia points, where a ridge terminates or separates into two ridges.

An orientation field shown in Fig. 2b of a fingerprint shown in Fig. 2a contains information about the local dominant orientations of fingerprint ridges, from which some other features can be derived such as singular points and dominant ridge line flow as shown in Fig. 2. The dominant ridge flow is represented by a set of curves running parallel to the ridges lines but not necessarily coinciding with ridges and valleys. There are two types of singular points: core and delta points. A core point is the turning point of an inner-most ridge and a delta point is a place where two ridges running side by side diverge. Orientation field, dominant ridge flow, and singular points are useful features for classification. A local ridge frequency is the number of ridges per unit length along a hypothetical segment orthogonal to the local ridge orientation. Its' inverse is the local ridge distance as shown in Fig. 2a. Although the local ridge distance varies across different fingers, it is difficult to serve as a reliable feature due to its high within-finger variation caused by the discontinuity of ridges and valleys and various unfavorable skin and imaging conditions. However, the average ridge distance over a fingerprint shows a stable and reliable feature and is employed in some approaches [12, 13].

Minutia points as shown in Fig. 2a are in general stable and robust to fingerprint impression conditions. They often serve as discriminative features for exact matching in most automatic fingerprint recognition systems. However, some fingerprint indexing approaches [9, 10] also use minutiae for coarse level fingerprint search. Another type of feature is the filter response of fingerprint image. Gabor filters are oriented band-pass filter with adjustable frequency, orientation, and bandwidth parameters. The responses of Gabor filters capture information of fingerprint local orientation, ridge frequency, and ridge discontinuity and hence can be used for both coarse level classification [5] and exact matching.



Fingerprint Classification. **Figure 2** A fingerprint image and its feature representation. The orientation field consisting of fingerprint local orientations is represented by short lines in (**b**). Core and delta points are marked in both (**a**) and (**b**) by *filled circles* and *triangles*, respectively. Two examples of ridge ending and ridge bifurcation, called minutia points, are enclosed by *circles* and *squares* in (**a**), respectively. An example of local ridge distance is shown by two *arrows* in (**a**). Three dominant *ridge flow curves* that can represent the Galton–Henry classes (here: right loop) are shown in both (**a**) and (**b**).

Classification Under Galton–Henry Scheme

Over the last four decades, many techniques have been developed for the automatic classification of fingerprints under Galton–Henry scheme, which can be coarsely assigned to one of these categories: rule-based, syntactic, structural, statistical and other approaches. While rule-based, syntactic and structural approaches are mainly used to partition the database into the human-interpretable fingerprint classes defined by Galton–Henry Scheme, statistical approaches are able to classify fingerprints into compact clusters in feature space.

The rule-based approaches codify the human expert knowledge of manual classification such as the singularity and the geometrical shape of ridge lines. It is not difficult to see from Figs. 1 and 2 that the five human-interpretable fingerprint classes can be determined by the number and location of the singular points plus some local ridge orientations. Fingerprints with neither core nor delta points are classified as arch. Whorls (plain whorl and twin loop whorl) have one or two cores and two deltas. Loops and tented arch contain only one core and one delta. Tented arch is discriminated from loops by examining the local orientations lying along the line connecting the core and delta points. The difference between these local orientations and the slope of the line is much smaller for a tented arch than loops. Left and right loops are distinguished by examining the local orientations around the core point with respect to the slope of the line [6]. Although a rule-based approach is simple and work well on rolled fingerprint with high image quality, robust and consistent detection of singular points in a poor quality fingerprint remains a difficult task. Thus, the rule-based approaches are in general sensitive to noise and cannot work on the partial fingerprint where the delta point is often missing.

A syntactic method represents a fingerprint by a sentence of a language extracted from the ridge flow or orientation field. For example, the three dominant ridge flow curves in Fig. 2 show the typical pattern of right loop. It is not difficult to see from Figs. 1 and 2 that, in general, the five human-interpretable fingerprint classes can be distinguished by such dominant ridge flow curves. In the syntactic approaches, a grammar is defined for each fingerprint class to build up sentences. Classification is performed by determining which grammar most likely generates the sentence extracted from a query fingerprint. In general, syntactic methods tend to be robust in the presence of image noise but often require very complex grammars to struggle against the large intra-class and small inter-class variations. Complex grammars often result in unstable classification.

The structural approaches organize low-level features into higher-level structure. One approach partitions the orientation field into connected regions characterized by homogeneous local orientations [11]. For example, it is not difficult to identify some homogeneous orientation regions from the orientation field shown in Fig. 2b. A relational graph that shows the relations among these regions of a fingerprint contains discriminative information for classification. An inexact graph matching technique is exploited to compare the relational graphs with class-prototypes. As a robust and consistent partition of orientation field is not an easy task, a template-based matching is developed to guide the partitioning [11]. Another approach converts the two-dimensional fingerprint structure into one-dimensional sequence and exploits hidden Markov model for classification [8]. A set of horizontal lines across the fingerprint is used to extract a sequence of features. It captures information about the local orientations and ridge distances and thus has higher discrimination power than the orientation field alone. Since the structural approaches rely on global structural information, they can work on noisy images and are able to deal with partial fingerprints where some singular points are not available.

Statistical approaches extract a fix-size numerical feature vector from a fingerprint and exploit statistical classifiers, such as *k*-nearest neighbor classifiers, support vector machines and artificial neural networks. The feature vector can be constructed based on the orientation field [4, 11, 12] or the responses of Gabor filters [5]. As features extracted from different fingerprint regions show different discriminating power, some weighting schemes [4, 11, 12] or non-uniform spacing techniques [5, 13] are developed to put higher weights in more discriminative regions of fingerprint. Karhunen–Loève (KL) transform and multi-space KL (MKL) transform [14] are also applied to reduce the dimensionality of feature vector. Statistical classifiers in general need to be trained with a fingerprint database. As Galton–Henry scheme defines the human-interpretable fingerprint classes rather than the natural

clusters of fingerprints in feature space, supervised training using fingerprint samples with known class labels is often applied. On the other hand, statistical approaches are able to classify fingerprints far beyond the Galton–Henry scheme into much more classes.

Classification with Machine-Generated Classes

The Galton–Henry Scheme does not offer much selectivity for fingerprint searching in large databases. Most automated systems [4–8] can only classify fingerprints into the five classes shown in Fig. 1 and the probabilities of the five classes are approximately 0.037, 0.029, 0.338, 0.317, and 0.279 for the arch, tented arch, left loop, right loop, and whorl, respectively [15]. The uneven distribution of these human-interpretable fingerprint classes further lowers the classification efficiency. In fact, for the application of the automated identification, it is often not obligatory to partition the database into human-interpretable fingerprint classes. Any classification scheme is in principle workable so long as different impressions of a same finger consistently fall into a same class. Instead of grouping fingerprints based on the visual appearance of fingerprint images, we can partition the database in the feature space into the machine-generated fingerprint classes, in the hope that more classes can be formed. However, there are always fingerprints located near the class boundaries regardless of how well the database is partitioned. These fingerprints are likely misclassified due to the large variations of different impressions of a same finger. To alleviate this problem, fingerprints are not pre-classified, but associated with numerical feature vectors. Given a query fingerprint, a fingerprint class is then formed by retrieving a portion of fingerprints from database whose feature vectors are similar or have small distance to that of the query fingerprint. Hence, this scheme is also called "continuous classification" [1, 11, 12].

Orientation field is often used to construct the numerical feature vector consisting of local orientations [4, 11–13]. Note that an orientation angle θ is a periodic variable with a period of 180° rather than 360° and has discontinuity at ±90° or 0° and 180°. The smallest and the largest angles in a period do not refer to two orientations far away but rather close to each other. The distance between two orientations θ^p and θ^q cannot be naively measured by $|\theta^p - \theta^q|$, but rather by $\min(|\theta^p - \theta^q|, 180^\circ - |\theta^p - \theta^q|)$. Thus, the distance between two feature vectors cannot be computed by simple arithmetic such as Euclidean distance. To simplify the distance computation, an orientation angle θ is decomposed into two component, $\cos(2\theta)$ and $\sin(2\theta)$ [1, 11, 14] so that the similarity of two fingerprints can be measured by the convenient dot product of the two feature vectors. This also enables to put weights on different orientations, for example, $r[\cos(2\theta), \sin(2\theta)]$, where r is the weight of orientation θ. In fact, the similarity of two feature vectors can be measured by the consistency of the orientation differences. Thus, a similarity measure between two feature vectors $\mathcal{O}^p = (\theta_1^p, \theta_2^p, ..., \theta_k^p, ...)$ and $\mathcal{O}^q = (\theta_1^q, \theta_2^q, ..., \theta_k^q, ...)$ is defined by $|\sum_k r_k \exp[2j(\theta_k^p - \theta_k^q)]|/\sum_k r_k$, where r_k are weights, $exp[\cdot]$ is a complex exponential function and $|\cdot|$ is a magnitude operator [12, 13]. Besides the orientation field, the average ridge distance over the fingerprint is also used as an auxiliary feature in some approaches [12, 13].

Given a query fingerprint, a fingerprint class is formed by retrieving a number of fingerprints from the database whose feature vectors are nearest to that of the query fingerprint. Depending on application scenarios, different fingerprint retrieval strategies can be applied, such as a fixed distance threshold, or a fixed percentage of fingerprints in database to be retrieved, or some combination of the both [12]. In an identification system, fingerprint retrieval and exact matching can be integrated so that the retrieval threshold increases from a small value until the query fingerprint is matched with one of the retrieved templates by a matching algorithm. The threshold can increase by a fixed step or based on a fixed number of newly retrieved fingerprints. The incorporation of matching in the fingerprint retrieval may greatly improve the retrieval performance if a good matching algorithm is applied [1, 11, 12].

The continuous classification in general needs to compare the feature vector of a query fingerprint with those of all fingerprints in the database. The time consumption of fingerprint retrieval thus directly depends on the database size. For large database, the continuous classification could be time consuming. To circumvent the one-by-one exhausting comparisons of a query fingerprint with all templates, database is partitioned into clusters and hence the query fingerprint

only needs to be compared with the cluster prototypes [12, 13]. Since in general there are always some fingerprints near the cluster boundaries regardless of how well the clusters are formed, it is crucial to retrieve, instead of one, a few clusters. For the application of automated identification, this clustering based classification scheme is comparable to the Galton–Henry scheme in terms of the search speed that is independent to the database size. But the former has potential to achieve better classification accuracy and efficiency. Fingerprint database indexing [9, 10] is a closely related problem to this classification scheme. Different from the clustering based classification scheme, however, fingerprint indexing approaches [9, 10] utilize minutia points that most automated fingerprint matching algorithms rely on for the exact fingerprint comparison.

Classification Performance

The performance of a fingerprint classification system is usually measured in terms of accuracy or error rate, efficiency or penetration rate, and speed or computational complexity. The measurements of these performance indicators could be quite different on different fingerprint databases. Therefore, the performance comparison of different classification algorithms should be based on the same database. The NIST (National Institute for Standards and Technology) Special Database 4 is the most often used database for the classification performance evaluation. It contains 2,000 fingerprint pairs, uniformly distributed in the five Galton–Henry classes (see Fig. 1). Some approaches are tested on a reduced set (called Set 2), containing 1,204 fingerprints extracted from the database according to the real distribution of fingerprints.

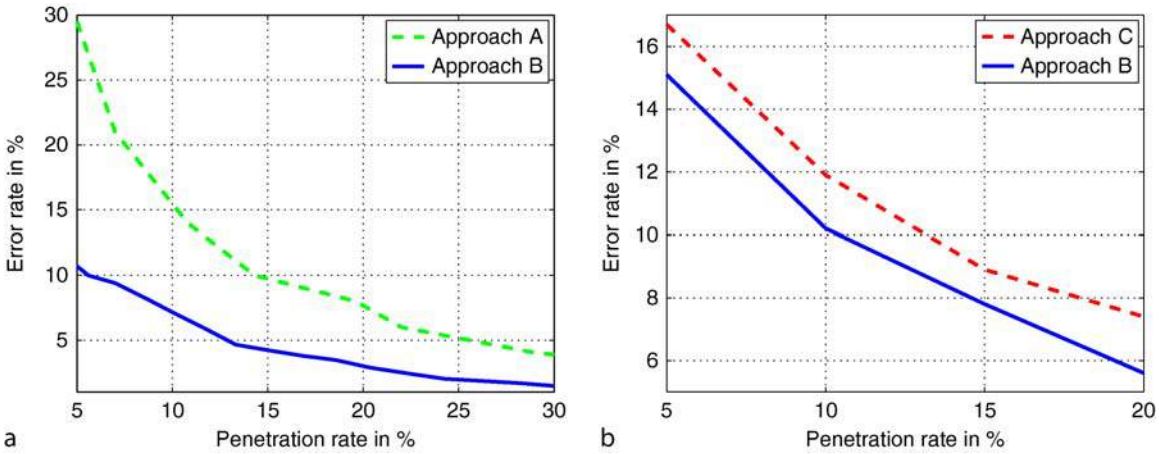
The error rate is computed as the ratio of the number of misclassified fingerprints to the total number of samples in the test set. For a Galton–Henry classification system, a fingerprint is misclassified if it is placed in a class different from the human assigned one as the true class membership of a fingerprint is determined by human experts. For a system that is based on the machine-generated fingerprint classes, a query fingerprint is misclassified if the retrieved subset from database contains no fingerprint originating from the same finger as that of the query fingerprint. The error rate of a classification system in general should be reported as a function of the penetration rate that is a performance indicator of the classification efficiency.

The classification efficiency is measured by the penetration rate defined as the average ratio of the number of fingerprints in a class to the total number of samples in the database [1, 11, 12]. If q_i represents the ratio of the number of fingerprints in class i to the total number of samples in database and p_i is the class occurrence probability, the penetration rate is calculated by $\sum_i p_i q_i$. For example, for the five Galton–Henry classes with the occurrence probabilities of 0.037, 0.029, 0.338, 0.317, and 0.279, respectively, the penetration rate of a error free classifier ($q_i = p_i$) is 0.2948, which lies between the penetration rates of 0.25 and 0.3333 for the four and three equal-sized classes, respectively.

Figure 3 illustrates the tradeoff between the classification error rate and the penetration rate of three techniques tested on two data sets. Obviously, lower classification error rate can be achieved at higher penetration rate. As higher classification accuracy and efficiency are measured by lower error rate and lower penetration rate, respectively, a lower curve indicates a better classification performance. Table 1 shows the classification results of some Galton–Henry scheme based approaches (the first seven rows) and the clustering based approach (the last two rows). All results are obtained from NIST Special Database 4. Some approaches are tested on the Set 2 and some approaches are tested on the second half of the database because they use the first half of the database to train their programs. Classification performance on the real distributed fingerprints is also resembled by the "weighted classes" shown in the third and the fifth columns. Note that Fig. 3 and Table 1 do not serve as a direct comparison between different algorithms due to different experimental settings and rate calculations. More information about the classification performances of these approaches can be found in the respective references [4–8, 10–12, 14].

Summary

The development of automatic fingerprint identification system for large database is a challenging task due to both accuracy and speed issues. Fingerprint classification as a tool to narrow down the searching space of



Fingerprint Classification. Figure 3 Classification error rate against penetration rate: (**a**) approach A in [11] and B in [12] tested on the the NIST Special Database 4 Set 2 containing 1,204 fingerprint pairs; (**b**) approach B in [12] and C in [10] tested on the the second half of the NIST Special Database 4 containing 1,000 fingerprint pairs.

Fingerprint Classification. Table 1 Classification error rates in % on NIST Special Database 4 of some Galton–Henry scheme based approaches (the first seven rows) and the clustering based approach (the last two rows)

Source	Five classes P.R. = 20%	Five weighted classes P.R. = 29.5%	Four classes P.R. = 28%	Four weighted classes P.R. = 29.7%	Test set
Candela et al. [4]	–	–	11.4	6.1	Second half
Karu and Jain [6]	14.6	11.9	8.6	9.4	Whole
Jain et al. [5]	10	7.0	5.2	–	Second half
Cappelli et al. [11]	–	12.9	–	–	Set 2
Cappelli et al. [14]	7.9	6.5	5.5	–	Second half
Senior [8]	–	–	–	5.1	Second half
Park and Park [7]	9.3	–	6.0	–	Whole
Jiang et al. [12]	5.3	3.3	3.5	3.2	Whole
Jiang et al. [12]	4.7	2.9	3.2	2.8	Set 2

The penetration rate is shown by the value of P.R. In the columns of "weighted classes", error rates of different classes are weighted by the class occurrence probabilities in the calculation of the total error rate

exact matching can alleviate these difficulties. A lot of different techniques have been developed to automate the Galton–Henry classification scheme, thanks to its human-interpretability and rigid segmentation of a database. However, the Galton–Henry classification scheme that partitions the database into human-interpretable fingerprint classes does not reduce the search space significantly. The database partition based on the machine-generated fingerprint classes seems to be a more promising alternative for efficient reduction of the search space. For a classification system that requires high accuracy, a fingerprint rejection engine can be applied to exclude poor quality fingerprints at a price of lower classification efficiency. Further research efforts are necessary to improve the classification performance.

Related Entries

► Fingerprint Features
► Fingerprint Indexing

▶ Fingerprint Recognition Overview
▶ Identification and Authentication

References

1. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer-Verlag, New York (2003)
2. Galton, F.: Finger Prints. McMillan, London (1892)
3. Henry, E.: Classification and Uses of Finger Prints. Routledge, London (1900)
4. Candela, G.T., Grother, P.J., Watson, C.I., Wilkinson, R.A., Wilson, C.L.: PCASYS – A pattern-level classification automation system for fingerprints. Technique Report: NIST TR 5647 (1995)
5. Jain, A.K., Prabhakar, S., Hong, L.: A multichannel approach to fingerprint classification. IEEE Trans. Pattern Anal. Mach. Intell. **21**(4), 348–359 (1999)
6. Karu, K., Jain, A.K.: Fingerprint classification. Pattern Recognit. **29**(3), 389–404 (1996)
7. Park, C.H., Park, H.: Fingerprint classification using fast fourier transform and nonlinear discriminant analysis. Pattern Recognit. **38**(4), 495–503 (2005)
8. Senior, A.: A combination fingerprint classification. IEEE Trans. Pattern Anal. Mach. Intell. **23**(10), 1165–1174 (2001)
9. Bhanu, B., Tan, X.: Fingerprint indexing based on novel features of minutiae triplets. IEEE Trans. Pattern Anal. Mach. Intell. **25**(5), 616–622 (2003)
10. Tan, X., Bhanu, B., Lin, Y.: Fingerprint identification: Classification vs. indexing. In: Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 151–156. Miami, Florida (2003)
11. Cappelli, R., Lumini, A., Maio, D., Maltoni, D.: Fingerprint classification by directional image partitioning. IEEE Trans. Pattern Anal. Mach. Intell. **21**(5), 402–421 (1999)
12. Jiang, X.D., Liu, M., Kot, A.: Fingerprint retrieval for identification. IEEE Trans. Inf. Forensics Secur. **1**(4), 532–542 (2006)
13. Liu, M., Jiang, X.D., Kot, A.: Efficient fingerprint search based on database clustering. Pattern Recognit. **40**(6), 1793–1803 (2007)
14. Cappelli, R., Maio, D., Maltoni, D.: Fingeprint classification based on multi-space KL. In: Proceedings of Workshop on Automatic Identification Advanced Technologies, pp. 117–120 (1999)
15. Wilson, C.L., Candela, G.T., Watson, C.I.: Neural network fingerprint classification. J. Artif. Neural Networks **1**(2), 203–228 (1993)

Fingerprint Comparing

▶ Fingerprint Matching, Automatic

Fingerprint Compression

Nigel M. Allinson
Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, UK

Synonym

Fingerprint Image Compression

Definition

Image files can be reduced in size by exploiting either more optimized data representation and not compromising the faithful recovery of the source image - lossless compression, or permitting recovery to within some distortion criteria - lossy compression. Fingerprint images are relatively large detailed images, and their compression can alleviate operational problems of transmission and storage.

Introduction

Fingerprint images, whether prints obtained directly from live subjects or forensically recovered latents, are normally recorded at 500 dots or pixels per inch (ppi) resolution with an 8-bit grayscale, though there is an increasing tendency to use a higher resolution of 1,000 ppi that permits accurate rendering of individual sweat pores along the ridge lines. A single digit print has a minimum area of about 20 mm $\times$ 15 mm, which yields a raw image of about 120 kB; while a tenprint record card (full set of individual slaps, rolls, and palm prints) requires several 10 **MB of storage. As national Automatic Fingerprint Identification Systems (AFISs) can contain tens of millions of individual record cards, storage requirements can be easily in excess of 100 TB [1]. It was the rapid rise in storage requirements for developing AFIS installations that drove the need for effective compression of reference fingerprint images. With the availability of low-cost, very high-capacity mass memory, this requirement may not be so clear today as it was in the early 1990s. Perhaps of greater importance is the need to transmit, over restricted bandwidth channels, both reference prints and latents recovered from crime and other scenes to remote

locations. Clearly, the transfer of images to a remote AFIS or between AFISs requires agreed standards for image compression that do not adversely affect the usage of fingerprints as a reliable and robust biometric.

Mainstream image compression aims to meet both the requirements of reducing storage requirements and enabling faster transmission. Significant compression is possible due to the limited acuity of the human visual system especially our low sensitivity in detecting low- contrast spatially-fine detail, the limited spatial resolution of electronic displays and some printing processes, and that humans are often tolerant of some degree of visible distortion. The normal image compression standards have been developed (1) to provide satisfactory reductions in memory requirements for all types of scenes, whereas fingerprints are a very restricted type of image; and (2) to provide an acceptable viewing experience for a viewer, whereas fingerprint image may be studied in great detail by an expert examiner and may be submitted to extensive automatic processing on AFIS systems.

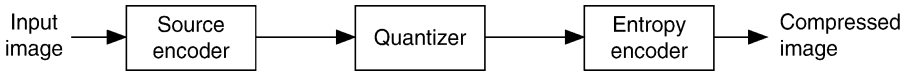
Fingerprint Image Standards

The typical ridge-valley period is approximately 500 μm with width of the ridge varying from about 100 - 300 μm. So a minimum image sampling frequency of 200–300 ppi would be sufficient to record unambiguously the friction skin details including local features or minutiae. FBI-compliant fingerprint scanners, and the resultant images, are specified at 500 ppi. The grey-scale resolution of some commercial fingerprint scanners is limited to only 2–3 bits though there is a general recognition that simple binary images, that is 1 bit deep, are unsatisfactory. For images to have the appearance of a continuous grey-scale approximately a minimum of 100 discrete levels are required, so current standards specify 8 bit or 1 byte deep grey- scale resolution (i.e., 256 possible levels). Sweat pores are smaller in diameter than a ridge width with an effective diameter that depends on whether the pore is open or closed. For the capture of pore structures, an image resolution of 1,000 ppi has been proposed [2] and is becoming to be accepted by the fingerprint community. These standards are embodied in current national and international standards for the data formats for the interchange of fingerprint information. These standards also recommend that the overall size of an image should range from 406 × 381 mm (800 × 750 pixels) for a single digit to 1,397 × 2,032 mm (2,750 × 4,000 pixels) for a complete palm print. This translates to image sizes ranging from about 0.5 MB to 10.5 MB for 500 ppi images and, of course, four times larger for 1,000 ppi images.

Image Compression

Image compression can be delineated into lossless and lossy coding schemes. The former referring to an encoding process that permits the original image to be retrieved without any degradation. Methods include Run-Length-Encoding where a consecutive row of three or more pixels with identical grey-scale or color values are represented by a two-byte pair. This forms part of several well-known image file standards such as TIFF (Tagged Image Format File) and PCX (PC Paintbrush Exchange). Another approach utilizes entropy coding which assigns codes to grey-scale or color values so that code lengths match with the inverse probabilities of these values. Reductions in storage requirements are usually very modest – typically less than 2:1 compression – and will not be discussed further. However, some AFIS installations do employ lossless compression for their archived reference images.

Lossy compression means that it is possible only to recover the original image to within some distortion criteria. The normal criterion for the acceptability is based on the non-visibility of coding artifacts under normal viewing conditions or, at least, the acceptability of these artifacts. For fingerprint images, the criteria need to include the absence of any artifacts that could subsequently interfere with future processing and feature recognition – either by AFIS algorithms or by a human expert. With inappropriate compression, it is more likely that legitimate queries will not be matched that is the FNMR or FRR, depending on the application, will increase. Several approaches have been explored for lossy compression but the dominant technique is based upon transforming an image from the spatial domain to a second domain which is based on spatial frequencies. Such an approach exploits the non-random distribution of the spatial frequencies in the localized objects that make up an image and the non-uniform sensitivity to differing frequencies by our visual system. The basic structure of a transform-based image coder is illustrated in Fig. 1. The decoder, which



Fingerprint Compression. **Figure 1** Overall structure of generic transform-based image coder.

recovers the best approximation of the original image from the transformed one, is essentially the reverse process. The source encoder employs a (usually) linear transform such as the Discrete Fourier Transform, Discrete Cosine Transform (DCT) or Discrete Wavelet Transform (DWT) to convert the entire image or some region of it into the different representational domain. The quantizer reduces the number of bits employed to represent the coefficients of the previous transform. This loss in precision, a many-to-few mapping, is the major source of compression in the overall encoding process. Quantization may be performed on individual transform coefficients, termed Scalar Quantization; or on a group of coefficients, termed Vector Quantization. As there are usually some correlation between consecutive coefficients which can be usefully exploited, vector quantization is generally more efficient than the scalar. Small valued coefficients, below a predefined threshold value, are ignored. Quantizers may uniformly reduce the precision of coefficients regardless of their magnitude or, more likely, they implement a non-uniform approximation by giving greater weight to higher valued coefficients. The final stage is an entropy encoder which losslessly compresses the quantized coefficients to yield a smaller output code stream. Typical methods include Huffman and Arithmetic coding with both being variable-length coding schemes but the former applied to individual coefficients while the latter is applied to a group of coefficients. Image compression possesses an extensive literature and a useful introduction is provided in [3].

The most common image compression is the JPEG (Joint Photographic Expert Group) standard which is DCT-based; and only the baseline encoder will be discussed here that sequentially compresses a stream of 8×8 pixel blocks of the image. Each block progresses through each processing step to yield a compressed output data stream. As adjacent image pixels are highly correlated, the forward DCT is the basis for achieving data compression by focusing most of the energy into the lower spatial frequency bands. The DCT causes no loss to the source image but simply transforms it into a domain where they can be efficiently encoded. Each of the 64 DCT coefficients is uniformly quantized according to a 64-element quantization table, which takes into account the falling sensitivity of the human eye to fine spatial details. After quantization, the quantized coefficients are ordered in a zig-zag sequence to assist the entropy encoding by placing low-frequency non-zero coefficients before high-frequency coefficients. The DC coefficient, which contains a significant fraction of the total image energy, is differently encoded. Decoding is essentially the reverse process. JPEG compression is efficient and simple to implement especially in dedicated hardware. Good compression rates can be achieved with little loss of perceived fidelity for naturalistic scenes up to 20:1 or 30:1 compression ratios. The use of 8×8 pixel blocks does at higher compression ratios create objectionable "blocking artifacts" especially in regions of low image contrast.

The basis function for JPEG is a discrete set of orthogonal cosine waves. Such sinusoidal waves are continuous in the spatial domain and are, in some sense, artificially truncated. There are numerous possible basis functions – some with limited support, that is they possess only a non-zero value for a limited interval. Wavelets are such a function which are defined over a limited distance and possess a zero average. From a single prototype wavelet function, the basis set is defined by a series of dilations and contractions of the prototype. Wavelets are a group of mathematical functions, of which the earliest example is the well-known Gabor function. These functions can be approximated as discrete filter structures. The variety of wavelet scales can be achieved efficiently using a cascade of high and low pass filters that decompose the image into several subbands, with each subband possessing optimal filter coefficients to match the image statistics for that band. Different numbers of subbands and their scope (bandwidth, orientation, etc), termed decomposition trees, are possible; and details of these, and other aspects of wavelet compression, are beyond the scope of this essay and the reader is referred to [4, 5]. As the wavelet transform is applied to the entire image and basis functions can overlap, there are no blocking artifacts. The type of artifact visible in highly compressed images is now low-level "ringing" around high-contrast edges. Wavelets, because of their local support, mimic the different

scales of receptive fields found in the human visual system and so produce visually more appealing images compared to JPEG images compressed to the same degree. JPEG, with its reliance on coding small blocks, is limited to moderate compression ratios; while DWT-based coding provides significant improvement in picture quality up to compression ratios of 70:1–100:1 for naturalistic scenes.

Fingerprint images are a very constrained class of image and as such it is reasonable to expect that more optimum forms of compression exist than provided by the standard methods. They are also exposed to more detailed scrutiny than most other images and they are subjected to extensive image processing and pattern recognition algorithms when submitted to an AFIS system. Though many quality metrics can be used to quantify the distortion introduced in lossy compression ranging from generic measures such as Peak-Signal-to- Noise (PSNR) to those developed specifically for fingerprints (such as the Image Quality Metric (IQM) [6]). PSNR, for 8-bit images, is defined as:

$$PSNR = 20\log_{10}\left(\frac{255}{e_{mse}}\right) \tag{1}$$

where the mean square error (e_{mse}) is given by

$$e_{mse} = \frac{1}{MN}\sum_{m-1}\sum_{n-1}[u(m,n) - v(m,n)]^2 \tag{2}$$

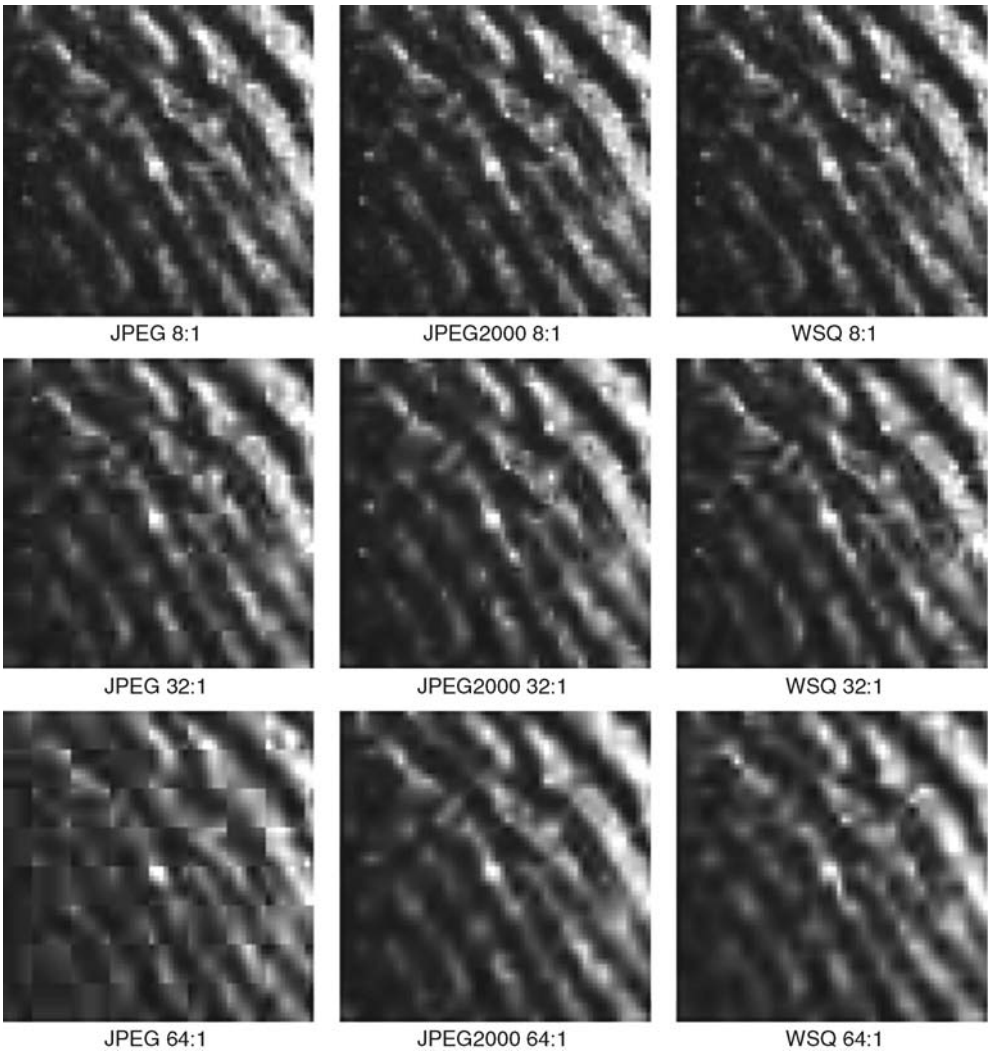
where $u(\bullet)$ and $v(\bullet)$ are the original and compressed images respectively – each of size M × N pixels. Higher PSNR means less distortion with no distortion equating to a PSNR = 48.13 dB. It is a useful metric in comparing similar image types. However, the "ground truth" for any compression scheme is the effect it has on the ability of fingerprint experts to make the same decisions as for the corresponding uncompressed image and an AFIS system to recover the correct match (or, when searching for matches to a latent, to rank consistently the most likely tenprint candidates).

Common Fingerprint Compression Methods

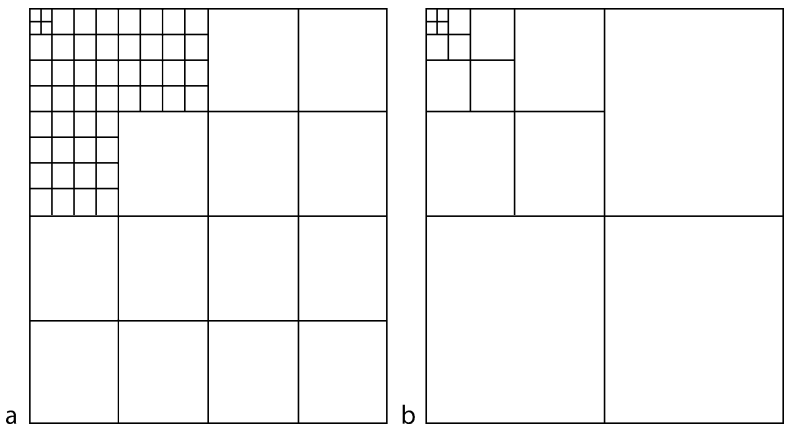
Three main compression methods have been applied for the storage, transmission and, display of fingerprints namely, JPEG, WSQ, and JPEG2000. WSQ (Wavelet Scalar Quantization) was developed, by the FBI in association with Los Alamos National Laboratory and NIST, specifically to reduce the media storage requirements of the FBI's expanding AFIS facility by providing lossy compression over the range 10:1 to 20:1. It has become an accepted standard for 500 ppi image storage and transmission. A set of typical set of compressed image for a latent fingerprint for these three methods is shown in Fig. 2.

Though the unsuitability of JPEG compression has been known for some time [7] as it suffers from visible blocking artifacts and loss of fine details (e.g., ridge pores) even at relatively low compression rates, it is still employed in some systems. Wide-area AFIS systems with connections to remote terminals often employ such JPEG compression to return reference tenprint images of potential matches to bureaux. For example, the UK national AFIS (*Ident 1*) displays images of potential tenprint matches as 12:1 JPEGs. It is possible for examiners to discern deterioration in such images.

JPEG2000 [8] is a relatively new standard for general-purpose image compression which attempts to address the limitations of JPEG as well as incorporating many other functions. It offers both lossless and lossy compression, provides a tiled representation of images at multiple resolutions, nonuniform compression to preserve greater detail in some region of interest, and embedded metadata and security functions within the image file. Both WSQ and JPEG2000 schemes are DWT-based, but with major differences in the form of the decomposition tree, quantization, and entropy coding employed. The WSQ uses the Daubechies (9,7) filter [9] to perform the DWT and the same filter is the default for the lossy JPEG2000 transform. The JPEG2000 employs a dyadic decomposition tree, while WSQ employs a fixed structure with 64 subbands (Fig. 3). The greater decomposition structure of WSQ may enhance compression as it approximates to orthnormalization and be better suited to the high spatial frequency content of fingerprints over more general imagery. The decomposition structure influences the number and length of the zero coding runs and so enhances compression, while the bit-plane scanning order of JPEG2000 permits finer control to achieve an arbitrarily specified compression rate. Both schemes use scalar quantization with JPEG2000 having the quantization step varying in response to the dynamic range of the respective subband. While for WSQ, all quantizer steps are uniform except for a



Fingerprint Compression. Figure 2 Comparative example of compressed image of 64 × 64 pixel region of latent.



Fingerprint Compression. Figure 3 Schematics of decomposition trees, (**a**) WSQ and (**b**) JPEG2000 (default).

lengthened middle interval. For the last coding step, WSQ employs Huffman entropy coding while JPEG2000 uses scalar arithmetic or trellis-coded quantization. Much of the flexibility and power of the JPEG2000 comes from the Embedded Block Coding with Optimized Truncation (EBCOT) algorithm. Wavelet coefficients from small blocks of the image are processed using the EBCOT algorithm which adapts the quantizer based on the statistics of the source image at the bit- plane level. These blocks are tiled with integrated header information concerning with coding details, and further quantization can be performed on the final bit-stream. The flexibility of this final stage in JPEG2000 encoding greatly assists in providing more optimal compression. WSQ employs a much simpler approach and coding details is calculated for each image and the coefficients included in the file header.

Fingerprint Compression Performance

WSQ has proved to be satisfactory at compressing fingerprint images by factors up to about 20, though the original requirement was to compress to 0.75 bpp (i.e., 10.7:1 compression). Watson and Wilson [10] report that experiments using WSQ compressed images with three different matching systems under the conditions that a FAR is 0.001 is maintained and the maximum reduction in the TAR of less than 0.01 is permitted, then there is little effect on performance for compression ratios less than 20:1. A few studies have compared the relative merits of WSQ and JPEG2000. Figueroa-Villanueva et al. [11] showed a significant improvement for JPEG2000 over WSQ at 0.75 bpp compression in terms of PSNR and Receiver Operating Curves (ROCs) for different sources namely, capacitive sensor, optical sensor, and scanned inked prints. A study of JPEG2000 and WSQ interoperability [12] concluded that JPEG2000 produced a slightly lower quality reconstructed image compared to WSQ for the same file size. Most studies have focused on coding high-quality inked prints or live print capture from various sources, and not on poorer quality latents. One study that involved latents and performance on an operational AFIS system [13] concluded that for compression ratios less than 32:1, JPEG2000 consistently produced higher identification rates than WSQ. There was also strong indication that moderate degrees of compression facilitated improved identification rates under normal operating procedures than uncompressed latent images.

Current national and international standards [14, 15] recommend that WSQ encoding is used for 500 ppi fingerprint images with compression limited to 15:1; but for images with resolutions greater than 500 ppi, 15:1 JPEG2000 should be employed. The UK national fingerprint system permits latents to be transmitted and submitted to *Ident 1* at 15:1 JPEG2000.

Conclusions

It may appear that satisfactory standards exist for compressing fingerprint images, certainly for the normal operational requirements associated with the effective transmission and storage of reference and livescan images. General enhancements in providing lower-cost, higher-capacity mass storage and increased bandwidth across both fixed and wireless data networks will reduce the pressure to develop new compression standards. However, fingerprints are an unusual and fairly well-defined class of image. Compression schemes such as JPEG2000 have been developed to cope well for a very wide variety of imagery and WSQ, though based on a principled consideration of the statistical properties of fingerprints, was developed prior to many more general advances in image compression. Recent proposals for improved fingerprint image compression are generally based on wavelet transformations but with more effective decomposition trees, optimized filter structures and coefficients through the use of genetic algorithm optimization, and vector quantization. The relationship between image enhancement and compression is not fully understood. There are suggestions that the filtering that occurs during compression may be advantageous in increasing identification, especially for latents. This is an avenue that needs to be explored further.

Related Entries

- ▶ Fingerprint Image Enhancement
- ▶ Fingerprint Matching, Automatic
- ▶ Fingerprint Matching, Manual
- ▶ Fingerprint Recognition, Overview

References

1. Komarinski, P.: Automated Fingerprint Identification Systems (AFIS). Elsevier Science Technology (2005)
2. Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST): Friction Ridge Digital Imaging Guidelines, 1 edn. (2001). URL http://www.theiai.org/guidelines/swgfast/
3. Ghanbari, M.: Image compression to advanced video coding. In: Standard Codecs, Telecomunications Series 49. Institution Electrical Engineers (2003)
4. Burrus, C., Guo, H., Gopinath, R.: Introduction to Wavelets and Wavelet Transforms: A Primer. Prentice Hall (1997)
5. Vetterli, M., Kovacevic, J.: Wavelets and subband coding. Prentice-Hall, Inc. (1995)
6. Nill, N.: Image quality evaluation - image quality of fingerprint (IQF). Tech. rep., The MITRE Corporation (2007)
7. Hopper, T., Preston, F.: Compression of grey-scale fingerprint images. In: Proceedings DCC '92. Data Compression Conference, pp. 309–318 (1992). DOI 10.1109/DCC.1992.227450
8. Skodras, A., Christopoulos, C., Ebrahimi, T.: The JPEG 2000 still image compression standard. IEEE Signal Process Mag **18**(5), 36–58 (2001)
9. Antonini, M., Barlaud, M., Mathieu, P., Daubechies, I.: Image coding using wavelet transform. IEEE Trans Image Process **1**(2), 205–220 (1992)
10. Watson, C., Wilson, C.: Effect of image size and compression on one-to-one fingerprint matching. Tech. rep., National Institute of Standards and Technology (2005)
11. Figueroa-Villanueva, M., Ratha, N., Bolle, R.: A comparative performance analysis of JPEG 2000 vs. WSQ for fingerprint image compression. In: Audio- and Video-Based Biometric Person Authentication, Lecture Notes in Computer Science, p. 1059. Springer Berlin/Heidelberg (2003)
12. Lepley, M.: JPEG 2000 and WSQ image compression interoperability. Tech. rep., The MITRE Corporation (2001)
13. Allinson, N.M., Sivarajah, J., Gledhill, I., Carling, M., Allinson, L.J.: Robust wireless transmission of compressed latent fingerprint images. **IEEE Transactions on Information Forensics and Security **2**(3), 331–340 (2007). DOI 10.1109/TIFS.2007.902684
14. American National Standards Institute: American national standard for information systems - data format for the interchange of fingerprint, facial, & other biometric information. URL http://fingerprint.nist.gov/standard/
15. International Organization for Standardization: Information technology - biometric data interchange formats - part 4: Finger image data. URL http://www.iso.org/

Fingerprint Contrast Enhancement

▶ Fingerprint Image Enhancement

Fingerprint Corpora

▶ Fingerprint Databases and Evaluation

Fingerprint Data Interchange Format

▶ Finger Data Interchange Format, Standardization

Fingerprint Databases and Evaluation

Fernando Alonso-Fernandez, Julian Fierrez
Biometric Recognition Group – ATVS,
Escuela Politecnica Superior,
Universidad Autonoma de Madrid,
Campus de Cantoblanco, Madrid 28049, Spain

Synonyms

Fingerprint benchmark; Fingerprint corpora

Definition

Fingerprint databases are structured collections of fingerprint data mainly used for either evaluation or operational recognition purposes.

The fingerprints in databases for evaluation are usually detached from the identity of the corresponding individuals, are publicly available for research purposes, and usually consist of raw fingerprint images acquired with live-scan sensors or digitized from inked fingerprint impressions on paper. These databases are the basis for research in automatic fingerprint recognition, and together with specific experimental protocols, are the basis for a number of technology evaluations and benchmarks. This is the type of fingerprint databases further developed here.

On the other hand, fingerprint databases for operational recognition are typically proprietary, usually incorporate personal information about the enrolled people together with the fingerprint data, and can incorporate either raw fingerprint image data or some form of distinctive fingerprint descriptors such as minutiae templates. These fingerprint databases represent one of the modules in operational automated fingerprint recognition systems, and will not be addressed here.

Fingerprint Databases for Evaluation

Among all biometric techniques, fingerprint recognition is the most widespread in personal identification due to its permanence and uniqueness [1]. Fingerprints are being increasingly used not only in forensic investigations, but also in a large number of convenience applications, such as access control or online identification [2].

The growth that the field has experienced over the past two decades has led to the appearance of increasing numbers of biometric databases for research and evaluation purposes, either ▶ monomodal (one biometric trait sensed) or ▶ multimodal (two or more biometric traits sensed). Previous to the databases acquired within the framework of the International Fingerprint Verification Competition series, the only large, publicly available datasets were the NIST databases [3]. However, these databases were not well suited for the evaluation of algorithms operating with live-scan images [1] and will not be described here. In this section, the authors present some of the most popular publicly available biometric databases, either monomodal or multimodal, that include the fingerprint trait acquired with ▶ live-scan sensors.

FVC Databases

Four international Fingerprint Verification Competitions (FVC) have been organized in 2000, 2002, 2004 and 2006 [4, 5, 6, 7]. For each competition, four databases were acquired using three different sensors and the SFinGE synthetic generator [1]. Each database has 110 fingers (150 in FVC2006) with eight impressions per finger (12 in FVC2006), resulting in 880 impressions (1,800 in FVC2006). In the four competitions, the SFinGe synthetic generator was tuned to simulate the main perturbations introduced in the acquisition of the three real databases.

1. In FVC2000 [4], the acquisition conditions were different for each database (e.g., interleaving/not interleaving the acquisition of different fingers, periodical cleaning/no cleaning of the sensor). For all the databases, no care was taken to assure a minimum quality of the fingerprints; in addition, a maximum rotation and a non-null overlapping area were assured for impressions from the same finger.
2. In FVC2002 [5], the acquisition conditions were the same for each database: interleaved acquisition of different fingers to maximize differences in finger placement, no care was taken in assuring a minimum quality of the fingerprints and the sensors were not periodically cleaned. During some sessions, individuals were asked to: (1) exaggerate displacement or rotation or, (2) have their fingers dried or moistened.
3. The FVC2004 databases [6] were collected with the aim of creating a more difficult benchmark because, in FVC2002, top algorithms achieved accuracies close to 100% [6]. Therefore, more intra-class variation was introduced. During the different sessions, individuals were asked to: (1) put the finger at slightly different vertical position, (2) apply low or high pressure against the sensor, (3) exaggerate skin distortion and rotation, and (4) have their fingers dried or moistened. No care was taken to assure a minimum quality of the fingerprints and the sensors were not periodically cleaned. Also, the acquisition of different fingers were interleaved to maximize differences in finger placement. Effects of quality degradation in fingerprint images can be observed in Fig. 1.
4. For the 2006 edition [7], no deliberate difficulties were introduced in the acquisition as it was done in the previous editions (such as exaggerated distortion, large amounts of rotation and displacement, wet/dry impressions, etc.), but the population was more heterogeneous, including manual workers and elderly people. Also, no constraints were enforced to guarantee a minimum quality in the acquired images and the final datasets were selected from a larger database (the BioSec multimodal database [8]) by choosing the most difficult fingers according to a quality index, to make the benchmark sufficiently difficult for an evaluation.

Fingerprint Databases and Evaluation. **Figure 1** Examples of quality degradation in fingerprint images due to factors like low/high pressure, dryness/moisture, dirt, etc.

BIOMET Multimodal Database

Five different biometric modalities are present in the BIOMET database [9]: audio, face image, hand image, fingerprint and signature. This database was designed with the additional goal of including unusual sensors (face images captured with an infrared camera and with a 3D acquisition system). The database consists of three different acquisition sessions. The number of individuals participating to the collection of the database was 130 for the first session, 106 for the second, and 91 for the last one, resulting in 91 individuals who completed the whole acquisition process. For fingerprint acquisition, an optical and a capacitive sensor were used. During the first acquisition campaign, only the optical sensor was used, whereas both the optical and capacitive sensors were employed for the second ant third campaigns. The total number of available fingerprints per sensor in the BIOMET database is 6 for the middle and index fingers of each contributor.

MCYT Bimodal Database

A large biometric database acquisition process was launched in 2001 by four Spanish academic institutions within the MCYT project [10]. The MCYT database includes ten-print acquisition (MCYT Fingerprint subcorpus) and on-line signature (MCYT Signature subcorpus) samples of each individual enrolled in the database. A total of 330 individuals were acquired in the four institutions participating in the MCYT project. Regarding the MCYT Fingerprint subcorpus, for each individual, 12 samples of each finger were acquired using an optical and a capacitive sensor under different control conditions. The MCYT database has been extended with the comprehensive BiosecurID multimodal database [11], which includes 8 different biometric traits from 400 donors collected in 4 sessions separated in time.

BioSec Multimodal Database

BioSec was an Integrated Project of the Sixth European Framework Programme which involved over 20 partners from nine European countries. The goal of BioSec was to leverage the integration of biometrics in a wide spectrum of everyday's applications. One of the activities within BioSec was the acquisition of a multimodal database. This database was acquired at four different European sites and includes face, speech, fingerprint and iris recordings. The baseline corpus [8] comprises 200 subjects with two acquisition sessions per subject. The extended version of the BioSec database comprises 250 subjects with four sessions per subject (about 1 month between sessions). Each subject provided in each session four samples of each of four fingers (left and right index and middle). Fingerprints were acquired using three different sensors. Some example images are shown in Fig. 2.

BioSecure Multimodal Database

The acquisition of the BioSecure Multimodal Database (BMDB) was jointly conducted by 11 European institutions participating in the BioSecure Network of Excellence. [11] The BMDB is comprised of three different datasets [12], namely:

1. *Data Set 1 (DS1)*, acquired over the Internet under unsupervised conditions (i.e., connecting to an

Fingerprint Databases and Evaluation. Figure 2 Example fingerprint images of two fingers acquired with three different sensors (from the BioSec baseline corpus). Fingerprint images of the same finger are shown for a capacitive sensor (*left of each subplot*), an optical sensor (*center*) and a thermal sensor (*right*).

URL and following the instructions provided on the screen).

2. *Data Set 2 (DS2)*, acquired in a standard office room environment using a PC and a number of commercial sensors under the guidance of a human supervisor.
3. *Data Set 3 (DS3)*, acquired using two mobile ▶ hand-held devices under two acquisition conditions (controlled-indoor and uncontrolled-outdoor).

The three datasets of the BMDB include a common part of audio and video data. Additionally, DS2 includes signature, fingerprint, hand and iris data, and DS3 includes signature and fingerprint data. The three datasets were acquired in two different sessions (approximately 2 months between them). Pending yet to be distributed publicly, the BioSecure multimodal database has approximately 1,000 subjects in DS1, and 700 in DS2 and DS3. Fingerprint data in DS2 were acquired using an optical and a capacitive sensor. Fingerprint data in DS3 were acquired with a PDA.

The databases MCYT, BiosecurID, BioSec, and BioSecure have some commonalities that enable their integration for specific research studies, e.g., on time variability and sensor interoperability [12].

Fingerprint Evaluation Campaigns

The most important evaluation campaigns carried out in the fingerprint modality are the NIST Fingerprint Vendor Technology Evaluation (FpVTE2003) [13] and the four Fingerprint Verification Competitions (FVC), which took place in 2000 [4], 2002 [5], 2004 [6] and 2006 [7]. A comparative summary between FVC2004, FVC2006 and FpVTE2003 is given Table 1. An important evaluation is also the NIST Minutiae Interoperability Exchange Test (MINEX) [14].

Fingerprint Verification Competitions (FVC)

The Fingerprint Verification Competitions were organized with the aim of determining the state of the art in fingerprint verification. These competitions have received great attention both from academic and commercial organizations, and several research groups have used the FVC datasets for their own experiments later on. The number of participants and algorithms evaluated has increased in each new edition of the FVC. Also, to increase the number of participants, anonymous participation was allowed in 2002, 2004 and 2006. Additionally, the FVC2004 and FVC2006 were subdivided into: (1) *open category* and (2) *light category*. The light category aimed at evaluating algorithms under low computational resources, limited memory usage and small template size.

For each FVC competition, four databases were acquired using three different sensors and the SFinGE synthetic generator [1]. The size of each database was set at 110 fingers with eight impressions per finger (150 fingers with 12 impressions per finger in FVC2006). A subset of each database (all the impressions from ten fingers) was made available to the participants prior to the competition for algorithm tuning. The impressions from the remaining fingers were used for testing. Once tuned, participants submitted their algorithms as executable files to the evaluators. The executable files were then tested at the evaluator's site and the test data were not released until

Fingerprint Databases and Evaluation. Table 1 Comparative summary between FVC2004, FVC2006 and FpVTE2003 (adapted from [6])

Evaluation	FVC 2004	FVC 2006	FpVTE 2003
Algorithms	Open category: 41 Light category: 26	Open category: 44 Light category: 26	Large scale test (LST): 13 Medium scale test (MST): 18 Small scale test (SST): 3
Population	Students	Heterogeneous (including manual workers and elderly people)	Operational data from a variety of U.S. Government sources
Fingerprint format	Flat impressions from low-cost scanners	Flat impressions from low-cost scanners	Mixed formats (flat, slap and rolled) from various sources (paper cards, scanners)
Perturbations	Deliberately exaggerated perturbations	Selection of the most difficult images according to a quality index	Intrinsic low quality fingers and/or non-cooperative users
Data collection	Acquired for this event	From the BioSec database	From existing U.S. Government sources
Database size	Four databases, each containing 880 fingerprints from 110 fingers	Four databases, each containing 1,800 fingerprints from 150 fingers	48,105 fingerprints from 25,309 subjects
Anonymous participation	Allowed	Allowed	Not allowed
Best average EER (over all the databases used)	2.07 % (Open category)	2.16 % (Open category)	0.2 % (MST, the closest to the FVC open category)

Fingerprint Databases and Evaluation. Table 2 Results in terms of equal error rate (EER) of the best performing algorithm in each of the four databases of the FVC competitions

Database	2000	2002	2004	2006
DB1 (%)	0.67	0.10	1.97	5.56
DB2 (%)	0.61	0.14	1.58	0.02
DB3 (%)	3.64	0.37	1.18	1.53
DB4 (%)	1.99	0.10	0.61	0.27
Average	1.73	0.19	2.07	2.16

the evaluation concluded. In order to benchmark the algorithms, the evaluation was divided into: (1) ▶ genuine attempts: each fingerprint image is compared to the remaining images of the same finger, and (2) ▶ impostor attempts: the first impression of each finger is compared to the first image of the remaining fingers. In both cases, symmetric matches were avoided.

In Table 2, results of the best performing algorithm in each FVC competition are shown. Data in the 2000 and 2002 editions were acquired without special restrictions and, as observed in Table 2, error rates decrease significantly from 2000 to 2002, demonstrating in some sense the maturity of fingerprint verification systems. However, in the 2004 and 2006 editions, it is observed that error rates increase with respect to the 2002 edition due to the deliberate difficulties introduced in the data, thus revealing that degradation of quality has a severe impact on the recognition rates [15].

NIST Fingerprint Vendor Technology Evaluation (FpVTE2003)

The NIST Fingerprint Vendor Technology Evaluation (FpVTE2003) [13] aimed at: (1) comparing systems on a variety of fingerprint data and identifying the most accurate systems; (2) measuring the accuracy of fingerprint matching, identification, and verification on actual operational fingerprint data; and (3) determining the effect of a variety of variables on matcher accuracy. Eighteen different companies competed in the FpVTE, and 34 systems were evaluated.

Three separate subtests were performed in the FpVTE2003: (1) the large-scale test (LST), (2) the medium-scale test (MST), and (3) the small-scale test (SST). SST and MST tested matching accuracy using individual fingerprints, whereas LST used sets of fingerprint images. The size and structure of each test were designed to optimize competing analysis objectives, available data, available resources, computational characteristics of the algorithms and the desire to include all qualified participants. In particular, the sizes of MST and LST were only determined after a great deal of analysis of a variety of issues. Designing a well-balanced test to accommodate heterogeneous system architectures was a significant challenge.

Data in the FpVTE2003 came from a variety of U.S. Government sources, including low quality fingers of low quality sources. 48,105 sets of flat slap or rolled fingerprint sets from 25,309 individuals were used, with a total of 393,370 fingerprint images. The systems that resulted in the best accuracy performed consistently well over a variety of image types and data sources. Also, the accuracy of these systems was considerably better than the rest of the systems. Further important conclusions drawn from the FpVTE2003 included: (1) the number of fingers used and the fingerprint quality had the largest effect on system accuracy; (2) accuracy on controlled data was significantly higher than accuracy on operational data; (3) some systems were highly sensitive to the sources or types of fingerprints; and (4) accuracy dropped as subject age at time of capture increased.

NIST Minutiae Interoperability Exchange Test (MINEX)

The purpose of the NIST Minutiae Interoperability Exchange Test (MINEX) [14] was to determine the feasibility of using minutiae data (rather than image data) as the interchange medium for fingerprint information between different fingerprint matching systems, and to quantify the verification accuracy changes when minutiae from dissimilar systems are used for matching fingerprints. ► Interoperability of templates is affected by the method used to encode minutiae and the matcher used to compare the templates. There are different schemes for defining the method of locating, extracting, formatting and matching the minutiae information from a fingerprint image [1]. In the MINEX evaluation, proprietary template formats were compared to the ANSI INCITS 378-2004 template standard.

The images used for this test came from a variety of sensors, and included both live-scanned and non live-scanned rolled and plain impression types. No latent fingerprint images were used. Participants submitting a system had to provide an algorithm capable of extracting and matching a minutiae template using both their proprietary minutiae format and the ANSI INCITS 378-2004 minutiae data format standard. The most relevant results of the MINEX evaluation are:

1. In general, proprietary templates lead to better recognition performance than the ANSI INCITS 378-2004 template.
2. Some template generators produce standard templates that are matched more accurately than others. Some matchers compare templates more accurately than others. The leading vendors in generation are not always the leaders in matching and vice-versa.
3. Authentication accuracy of some matchers can be improved by replacing the vendors template generator with that from another vendor.
4. Performance is sensitive to the quality of the dataset. This applies to both proprietary and interoperable templates. Higher quality datasets provide reasonable interoperability, whereas lower quality datasets do not.

Related Entries

► Biometric Sample Acquisition
► Fingerprint Device
► Interoperability
► Performance
► Performance Evaluation

References

1. Maltoni, D., Maio, D., Jain, A., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003)
2. Jain, A., Ross, A., Pankanti, S.: Biometrics: A tool for information security. IEEE Trans. Inf. Forensics Secur. **1**, 125–143 (2002)
3. NIST Special Databases and Software from the Image Group, http://www.itl.nist.gov/iad/894.03/databases/defs/dbases.html
4. Maio, D., Maltoni, D., Capelli, R., Wayman, J., Jain, A.: FVC2000: Fingerprint verification competition. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 402–412 (2002)
5. Maio, D., Maltoni, D., Capelli, R., Wayman, J., Jain, A.: FVC2002: Second fingerprint verification competition. Proc. Intl. Conf. Pattern Recognit. **3**, 811–814 (2002)
6. Cappelli, R., Maio, D., Maltoni, D., Wayman, J.L., Jain, A.K.: Performance evaluation of fingerprint verification systems. IEEE Trans. Pattern Anal. Mach. Intell. **28**, 3–18 (2006)
7. FVC2006. Fingerprint Verification Competition, http://bias.csr.unibo.it/fvc2006/default.asp (2006)
8. Fierrez, J., Ortega-Garcia, J., Torre-Toledano, D., Gonzalez-Rodriguez, J.: BioSec baseline corpus: a multimodal biometric database. Pattern Recognit. **40,** 1389–1392 (2007)
9. Garcia-Salicetti, S., Beumier, C., Chollet, G., Dorizzi, B., les Jardins, J., Lunter, J., Ni, Y., Petrovska-Delacretaz, D.: BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. Lecture Notes Comput. Sci. **2688**, 845–853 (2003)
10. Ortega-Garcia, J., Fierrez-Aguilar, J., Simon, D., Gonzalez, J., Faundez-Zanuy, M., Espinosa, V., Satue, A., Hernaez, I., Igarza, J., Vivaracho, C., Escudero, D., Moro, Q.: MCYT baseline corpus: a bimodal biometric database. IEE Proc. Vision Image Signal Process **150**, 395–401 (2003)
11. Fierrez, J., Galbally, J., Ortega-Garcia, J., et al: BiosecurID: A multimodal biometric database. Pattern Anal. Appl. **12** (2009)
12. Ortega-Garcia, J., Fierrez, J., Alonso-Fernandez, F., et al: The multi-scenario multi-environment BioSecure multimodal database (BMDB), IEEE Trans. Pattern Anal. Mach. Intell. **31** (2009)
13. Wilson, C., et al: Fingerprint Vendor Techonology Evaluation 2003: Summary of Results and Analysis Report. NISTIR 7123, http://fpvte.nist.gov (2004)
14. Grother, P., et al.: MINEX – Performance and interoperability of the INCITS 378 fingerprint template, NISTIR 7296, http://fingerprint.nist.gov/minex (2005)
15. ANSI-INCITS 378, Fingerprint Minutiae Format for Data Interchange, American National Standard (2004)

Fingerprint Device

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Fingerprint Encryption

► Fingerprints Hashing

Fingerprint Fake Detection

Jean-François Mainguet
Grenoble, France

Synonyms

Liveness detection; Cut finger problem; Dead finger detection; Fake finger detection; Gummy bear finger; Latex finger; Liveness detection

Definition

Fingerprint fake detection is used to identify a fake finger, such as a cast made of latex. By extension, it also includes tests to detect a cut finger or dead finger, or a latent print remaining on a sensor after usage.

Introduction

In "Diamonds are forever" (1971) [1] James Bond took the identity of Peter Frank with a thin layer of latex glued on his fingertip to spoof Tiffany Case's camera. James was using a simple fake finger, but the situation can be worse. With automated fingerprint recognition systems becoming more widely used, concerns over fingerprint fake detection have increased. In March 2005, a team of carjackers in Subang Jaya in Malaysia chopped off part of the owner's left index finger, when they realized that his S-Class Mercedes Benz had a security feature which would immobilize the car without his fingerprint. Even with more reliable cut finger detectors in use, it is likely that this will happen again.

Security of a fingerprint-based system can be divided into two main areas:

1. The electronic security, which poses the question: "Is the electronic system, at the other end of the wires, a real trustful authorized fingerprint system?"

2. The liveness security, which asks a different question: "Is the object touching the sensor a real finger, alive and connected to a living person?"

Answers for electronic security deal with cryptography, using challenge-response schemes and cryptographic codes. Since the focus of this essay is to answer the second question, we will suppose that the electronic system is perfect and cannot be broken.

To begin, we know 100% security does not exist. However, what we would verify is that, "I'm Mr X, a living person not under threat and I agree to this action." Lacking the ability to read a person's mind, this is an impossible task. At the opposite end, a basic fingerprint system will identify a particular fingerprint image as likely the same one as registered in the template, which is only a small brick within a full security system.

To fill the gap, we need to acquire more information that will enable us to say "this is a real alive finger." If we can do that, then we have a good chance to know that a real person is making the transaction, rather a cast or cut finger being applied to the sensor. This will not answer the problem of detecting a person under threat, but it should be enough under normal usage, although some situations will never be detectable. For example, it will be impossible to detect a graft. In France, a man received two hands from a donor, a great medical achievement [2]. But at the same time, he received 10 brand new fingerprints! There is also the case of George who attempted to enter the US illegally on 24 September 2005 through the Nogales, Arizona Port of Entry during which time US Customs and Border Protection officers noted that his fingerprints had been surgically replaced with skin from his feet. George stated that this procedure had been done by a doctor in Phoenix to "clean" his identity [3]. But these should be extremely rare cases. What is primarily desired is to avoid anyone stealing a fingerprint to impersonate someone else. So, while it is impossible to create an absolute fake finger detection system, it is possible to make things extremely hard to be cracked.

Compromised Fingerprint

When someone creates a fake of one of the fingerprints and use it to spoof a fingerprint system, then we say that this fingerprint is compromised. With a smart card (or a key), the smart card can be revoked. Further use of the card can be prevented and a new one can be created. But with fingerprints, this is limited to the 10 fingers. Biometric traits – the basis of biometrics cannot be revoked.

Liveness detection solves the compromised fingerprint problem. If the system can check that it is the real alive finger, then there is no possibility of using a fake.

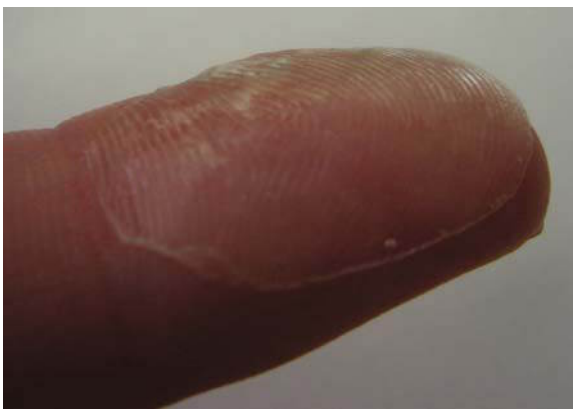
Attended/Unattended System

It is commonly admitted that an attended biometric system does not need any liveness detection because the supervisor "obviously" checks that a real alive person is present. In the case of fingerprints, this would be true if the supervisor was checking the finger: is the finger really connected to the body, and without any glued cast (Fig. 1)?

Fingerprint Fake Finger Detection Levels

There are three fake finger methods and detection levels described, starting from the easiest to the hardest to detect:

1. Latent print left on the sensor
2. Fake/copies:
   a. Printed fingerprint image
   b. Fake made of gelatin, latex, or other material



Fingerprint Fake Detection. Figure 1 Thin fake made of gelatine glued on a real finger.

c. Thin layer of material glued to a real finger, including real skin cells grown in a laboratory
3. Original finger:
   a. Cut out
   b. Belonging to a dead person
   c. Alive person under threat

Significant Developments in Fingerprint Spoofing

In the early 1990s, Ton van der Putte developed and improved a technique to fool the available biometrical fingerprint recognition systems. But when he contacted the manufacturers and showed them the security breach in their systems, it was ruled unimportant and nothing was done to solve it. In 2000, van der Putte and Jeroen Keuning decided to raise people's awareness and published an article [4] "*as a warning to those thinking of using new methods of identification without first examining the technical opportunities for compromising the identification mechanism.*" Using duplication with and without cooperation and material such as silicone rubber, van der Putte and Keuning fooled four optical sensors and two silicon-based capacitance sensors.

In 2001, Kàkona [5] described how to spoof an optical fingerprint sensor using a printed fingerprint (Fig. 2) and reactivating latent fingerprints on the sensor's surface by breathing on it. In 2002, Thalheim et al. [6] tested five solid-state and two optical fingerprint sensors. Gummy bears were introduced by Matsumoto [7] in 2002. The experiments involved 11 commercially available fingerprint sensors, both optical and capacitive, using a new cheap material, gelatin.

Further studies from Kang [8] and Blommè [9] extended the previous work. Stén et al. [10] spoofed a capacitance sensor using hot glue for the negative mold and gelatin for the fake finger. Marie Sandström (2004) published her thesis [11], "Liveness Detection in Fingerprint Recognition Systems," which gathered most of the available technologies at that time as well as experiment results on various sensors.

In 2006, Ton van der Putte updated his work [12] using additional material including silicon gel, acrylic paint, gelatin, gum arabic. Ongoing attempts to spoof fingerprint sensors continue to appear on the Internet; for instance, the Chaos Computer Club [13] used wood glue and published their results online (Fig. 3).

Fingerprint Fake Detection. **Figure 2** Printed fingerprint spoofing an optical sensor.

Fingerprint Fake Detection. Figure 3 Wood glue using a printed fingerprint as negative.

Making a Fake Fingerprint

Making a fake fingerprint always requires a fingerprint image. The easiest way to get a good fingerprint image is to have the cooperation of the donor. This is rarely the case in the real world, except when the latent print is left on the sensor (Level 1). In that case, the donor completes a successful acquisition; later, the impostor "reactivates" the latent print by breathing on the sensor. This happened in the past with some optical systems and with some capacitance-based sensors. A simple algorithm rejecting an image previously acquired is generally enough to avoid this problem, while swipe-sensing just eliminates this possibility.

The required fingerprint image is not necessarily exactly the same as the original fingerprint of the donor. Minutia matching (which is the main matching technology for fingerprints) only requires having the minutia locations and directions at the right place. It is possible, in theory, to create a fingerprint image with the right minutia locations that does not look like the original. This requires accessing the minutiae locations stored in the template, which should be ciphered. Work related to some form of automated reconstruction has been proposed, requiring only access to the matching score (hill-climbing) [14, 15]. This technique is far more difficult compared to obtaining the original fingerprint.

So in general, an impostor will take the easiest way to obtain the original fingerprint image. We will not deal here with the Level 3 which requires the original finger, cut, or belonging to a dead person. Obtaining the original image can be done with or without cooperation. With cooperation is the easiest way, and most

articles dealing with spoofing assume that the right finger is available to create a negative mould. Without cooperation will be the most common situation. Fortunately, stealing the fingerprint of someone else is not easy. Even for forensic professionals, it is hard to identify people from fingerprints left in a crime scene. Moreover, it is very difficult to select which fingerprint to use. It is likely that the forefinger is the most common finger used in a fingerprint system, but selecting the right fingerprint is not obvious.

Once the right image is obtained, image processing skills are generally required to enhance the fingerprint. Printed circuit technologies are often proposed to create a negative mould, but sometimes direct molding techniques, such as a rubber stamp (Fig. 4), can be used to get a positive.

With a negative mold, you need to create the positive cast that will be used to spoof the fingerprint sensor. Glue, latex, gelatin, and other materials have been proposed (Fig. 5), but the most difficult thing is to select the right material that properly fits the sensor. Latex may work for some sensors and not for others. Understanding the physics of the sensing techniques will help. So, at the end of the day, making a fake finger without cooperation is difficult, but far from being impossible.

Fingerprint Fake Detection. **Figure 4** Rubber stamp.

Liveness Measurement

To be able to detect a fake, we must first answer the question of what defines a live finger. Some activities related to liveness are:

1. Cellular metabolism with material transformation (protein)
2. Movement
3. Heat production (a sub-product)
4. Blood circulation for material delivery and heat transportation (regulation)

These activities have a number of signatures: physical, chemical, mechanical, nervous, geometrical, to name a few. Moreover, signification changes with the observation scale.

Detection methods can be active or passive. Active techniques involve a response to a stimulus, and can be voluntary or involuntary. It could be seen like a challenge-response as used in regular cryptographic techniques. Involuntary are reflexive challenge responses (removing your finger when you feel an electrical shock), while voluntary are behavioral challenge responses (how many vibrations did you feel?). Active detection is very interesting, because the nervous system up to the brain can be involved, which is a good marker of aliveness. But generally, active detection is not very practical from a user point of view, and nociceptive methods are not acceptable.

Passive techniques are linked to physiological activity of the finger. Here are some physiological data about fingers:

1. Cells, a bone, and a nail make a structure of about 1–10 cm^3. Note that there is no muscle (and so electrical activity is coming from other areas)
2. Arterial blood brings all chemicals, oxygen, and heat and returns to the body through veins
3. Skin is composed of three layers:
   a. Stratum corneum made of dead cells, more or less hydrated, 100 μm thick, variable electrical conductivity
   b. Blood-free epidermis, 0.05–1 mm thick, made of proteins, lipids, melanin-forming cells
   c. Dermis: dense connective tissues, capillaries arranged in vertical loops
4. Arteriovenous anastomoses, innervated by nerve fibers that regulate the blood flow of a factor of 30 in response to heat

Fingerprint Fake Detection. **Figure 5** Some moulded fakes: gelatine, plastic (negative), alginate, silicon.

5. Temperature range: 10°–40°C; not regulated
6. Skin emits some specific molecules (odor)
7. Skin presents some plasticity

Remark: The external layer of the skin is made of *dead* cells, which is not a favorable configuration for liveness detection!

Any liveness detection reader should read one or several data related to the previous list. Also, reading only one characteristic will not ensure that the read fingerprint is coming from a real finger: some material exhibiting the same plasticity than skin exists for instance.

Fingerprint Sensors with Liveness Detection

Few fingerprint sensor manufacturers claim to have some kind of liveness detection; and whenever claimed, little or evasive information is given. But, new techniques and ideas are being explored:

1. Maybe the most common liveness detection method is based on electrical measurements, using the conductivity and/or impedance of the skin. Some sensors can acquire fingerprints using electrical properties of the skin (RF-field, capacitance, electro-optical), and so require a conductive material to be spoofed. Non-conductive latex cannot work
2. Light transmission properties of the skin and/or the blood. Hospitals are using pulse oxymetry to measure the blood oxygenation, i.e., the percentage of oxyhemoglobin compared to deoxyhemoglobin. Two LEDs send infrared light through the finger to a photodiode, so it is some additional material aside the regular fingerprint sensor. Skin spectrum has also been proposed [16], using a wider range of colors
3. Perspiration induces detectable changes in time when looking at a series of images [17]
4. Distortion of the skin depends a lot on its plasticity [18]
5. Skin emits some specific molecules that can be detected (odor) [19]

Faking the Counter Measures

Any measurement can be faked:

1. Electrical method can be faked by the appropriate voltage applied on the sensing area (or even a simple connection to real skin while a fake is applied)
2. Optical methods can be faked by the appropriate plastic with the correct absorption characteristics
3. An optical sensor is made of photodiodes; it is always possible to send the appropriate light, synchronized with the light sent by the system
4. Cardiac pulse can be faked with the appropriate pump and pipes

But it is possible to make things very hard to spoof. For instance, the latest immigration control systems acquire the two forefingers at the same time, and so trying to spoof both sensors at the same time will be much harder.

Conclusion

Fake fingerprint detection will be an important feature of fingerprint sensors in the future, likely mandatory. We already know that a no fingerprint system will be 100% spoof-proofed, but several different sensors reading different information at the same time will be very hard to deceive. The "Swiss cheese" model applies here: each slice of cheese is not 100% secure, some holes exist. But more slices will stop most of threats. . . at the cost of each slice!

Related Entries

- ▶ Anatomy of Fingerprint
- ▶ Forensic Science
- ▶ Security

References

1. Danjaq, S.A.: DIAMONDS ARE FOREVER Copyright © (1971)
2. Transplantation des deux mains dans le service de chirurgie de transplantation de l'Hôpital Edouard Herriot/Hôpitaux de Lyon. January 13, 2000 http://www.chu-lyon.fr/internet/relations_medias/2005/5ans_double_greffe/dossier_presse_5ans_double_greffe.pdf
3. US Department of Justice's US Attorney's office for Arizona press release, May 3 (2006)
4. van der Putte, T., Keuning, J.: Biometrical fingerprint recognition don't get your fingers burned. In: Proceedings of IFIP TC8/WG8.8 Fourth Working Conference on Smart Card Research and Advanced Applications, pp. 289–303. Kluwer, Dordecht (2000) (http://cryptome.org/fake-prints.htm)
5. Kàkona, M.: Biometrics: yes or no? (http://home.i.cz/kakl/biometrics/Biometricsyesorno.htm). Accessed Feb 11, (2009)
6. Thalheim, L., Krissler, J., Ziegler, P.M.: Body check: biometric access protection devices and their programs put to the test (http://www.heise.de/ct/02/11/114/)
7. Matsumoto, T., Matsumoto, H., Yamada, K., Hoshino, S.: Impact of artificial gummy fingers on fingerprint systems. In: Proceedings of SPIE on Optical Security and Counterfeit Deterrence Techniques IV, San Jose, CA, vol. #4677 (2002) (http://cryptome.org/gummy.htm)
8. Kang, H., Lee, B., Kim, H., Shin D., Kim, J.: A study on performance evaluation of the liveness detection for various fingerprint sensor modules. http://www.springerlink.com/content/0df29gectgdkwrkl/
9. Blommé, J.: Evaluation of biometric security systems against artificial fingers (http://www.ep.liu.se/exjobb/isy/2003/3514/exjobb.pdf). Accessed Feb 11, (2009)
10. Stén, A., Kaseva, A., Virtanen, T.: Fooling fingerprint scanners – biometric vulnerabilities of the precise biometrics 100 SC scanner. In: Proceedings of the Fourth Australian Information Warfare and IT Security Conference, Adelaide, Australia (2003) http://www.stdot.com/pub/ffs_article_asten_akaseva.pdf
11. Sandström, M.: Liveness detection in fingerprint recognition systems. Ph.D. thesis, Linkoping University, Sweden (2004)
12. van der Putte, T.: Workshop spoofing fingerprints, SAFE-NL, University of Twente http://wwwes.cs.utwente.nl/safe-nl/meetings/08-06-2006/ton.pdf
13. Chaos Computer Club, How to fake fingerprints? October 26, 2004 http://www.ccc.de/biometrie/fingerabdruck_kopieren.xml?language = en
14. Martinez-Diaz, M., Fierrez-Aguilar, J., Alonso-Fernandez, F., Ortega-Garcia, J., Siguenza, J.A.: Hill-climbing and brute-force attacks on biometric systems: a case study in match-on-card fingerprint verification. In: Proceedings of the 40th Annual IEEE International Carnahan conferences security technology, Lexington, Kentucky, pp. 151–159 (2006)
15. Ross, A., Shah, J., Jain, A.K.: Towards reconstructing fingerprints from minutiae points. In: Proceedings of SPIE Conference on Biometric Technology for Human Identification II, Orlando, FL, pp. 68–80 (2005)
16. Nixon, K.A., Rowe, R.K.: Multispectral fingerprint imaging for spoof detection. In: Jain, A., Ratha, N. (eds.) Proceedings of SPIE on Biometric Technology for Human Identification II, Bellingham, WA, vol. 5779, pp. 214–225 (2005)
17. Parthasaradhi, S., Derakhshani, R., Hornak, L., Schuckers, SAC.: Time-series detection of perspiration as a liveness test in fingerprint devices. IEEE Trans. Syst. Man Cybern. C Appl. Rev. **35**, 335–343 (2005)

18. Antonelli, A., Cappelli, R., Maio, D., Maltoni, D.: A new approach to fake finger detection based on skin distortion. International Conference on Biometric Authentication (ICBA'06), Hong Kong, China (2006)
19. Baldisserra, D., Franco, A., Maio, D., Maltoni, D.: Fake fingerprint detection by odor analysis. International Conference on Biometric Authentication (ICBA'06), Hong Kong, China (2006)

Fingerprint Features

Josef Bigun
Embedded Intelligent Systems Center
Halmstad University, IDE, Halmstad, Sweden

Synonyms

Fingerprint analysis; Fingerprint characteristics; Fingerprint signatures

Definition

Fingerprint features are parameters in epidermis images of a fingertip (the fingerprint) that can be utilized to extract information which is exclusively specific to a unique person. These parameters can be measured by computational techniques applied to a digital image obtained by a *fingerprint sensing* method, e.g., using live optical or solid-state scanners, and digitizing ink-rolled or latent fingerprint images. Such identity characterizing parameters include one or more specifics of ridge–valley direction and frequency, *minutiae*, and *singular points*. The fingerprint features should be reproducible and resilient to variation in the face of external factors such as aging, scars, wear, humidity, and method of collection.

Introduction

Fingerprints consist of ridges alternating with valleys that mostly run in parallel but also change direction smoothly or may terminate abruptly. Other patterns in nature that resemble fingerprints include Zebra skins, corals, and shallow sea-bottom. Such pattern variations can be parametrized and used to characterize the fingerprints of individuals and to distinguish them from others. Identity establishment by fingerprint matching has been used by human experts long before the computer era, e.g., the nineteenth century contributors to the advancement of fingerprints, Jan. Purkyně, William Herschel, Alphonse Bertillon, Francis Galton, Edward Henry, Aziz-ul Haque, Chandra Bose, to name but few [1].

Caused by a foray of factors, low contrast and noisy images can compromise the reproducibility of fingerprint feature severely. Although the reason of poor image quality might be known, a better data acquisition is sometimes not a practicable option, e.g., latent fingerprints at a crime-scene, aging, scars and bruises, professional wear, etc. Accordingly, reproducibility is an important property of fingerprint features to be used. Another issue is their computational efficiency, if lacking it can hinder a practice of a fingerprint recognition method altogether, e.g., AFIS systems used in border-control, altogether.

Minutiae, to be discussed below in further details, represent the most widely used feature type by machine as well as human experts to determine if two fingerprints match. The geometric interrelationships of extracted minutiae, the spatial frequency between them or in their vicinity, and the local direction, contribute all to the strength of a minutiae based feature set so as to uniquely characterize a fingerprint. Another set of well-localized points is singular points. As will be detailed later, these are few, and one need larger neighborhoods to determine them in comparison to minutiae.

An important tool to characterize fingerprints is their *direction fields* since they are used in many operations of fingerprint processing. In the coming sections, we discuss direction field estimation, followed by minutiae, and singular points.

Direction Fields

The fingerprint direction fields are dense vector fields representing dominant local directions. A direction of an image-point (a pixel) is thus a property of its neighborhood; by itself no pixel can define a direction. Early direction fields were associated with local edges or lines and were approximated by the gradient of the image, $\nabla f = (\partial f/\partial x, \partial f/\partial y)^T$ where f is the local image, on digital lattices. Direction in this sense is the angle of the gradient and has already been used in

1960s, including in fingerprint applications. However, this concept hinders the use of effective signal processing tools, because a sinusoidal wave pattern (the local fingerprint) has a unique direction whereas half of its gradient directions in a fingerprint patch differ with 180° from the other half, resulting in a neither unique nor continuous representation if gradient angles would have defined feature spaces representing ridge directions. In turn this hinders efficient signal processing and inference which require rotation, scale-space, and interpolaton operations.

Direction Fields by Structure Tensor

An effective cure to representation ambiguity of ridge and valley direction is to use the concept of iso-curve (points having the same gray-value), which suggests the use of 2×2 tensors naturally, in the quest of an optimal direction estimation in the total least squares sense. This is summarized next, where the notion of image refers to a local patch of a fingerprint.

If all iso-curves of an image has a common direction the image is said to be linearly symmetric, e.g., sinusoidal planar waves resembling most neighborhoods of fingerprints. Ideally, the unknown direction $\mathbf{k}$ is optimal for an image $f(\mathbf{r})$ if the image is invariant to a translation in the amount of ε along the line $\mathbf{k}$ where ε is small and can be positive as well as negative, and $||\mathbf{k}||=1$. Then the total translation error $\mathcal{E}$

$$\begin{aligned}\mathcal{E}(\mathbf{r}) &= f(\mathbf{r}+\epsilon\mathbf{k}) - f(\mathbf{r}) \\ &= \epsilon[\nabla f(\mathbf{r})]^T\mathbf{k} + \mathcal{O}(\epsilon^2) = \epsilon e(\mathbf{r}) = \mathbf{0}\end{aligned} \tag{1}$$

will be zero for all $\mathbf{r}$ if the gray-value patch f is translation invariant in the direction $\mathbf{k}$. Here $e(\mathbf{r})$ is the unit-error. Ignoring the quadratic term $\mathcal{O}(\epsilon^2)$, because ε represents small translations, if and only if the unit-error of translation in the (fixed) direction $\mathbf{k}$

$$e(\mathbf{r}) = [\nabla f(\mathbf{r})]^T\mathbf{k} = \mathbf{0} \tag{2}$$

vanishes, (1) will vanish for *all* $\mathbf{r}$ of the patch. Evidently, the unit-error will even vanish on a discrete sub-set of the points of the patch, as below

$$\begin{pmatrix} D_x f_1 & D_y f_1 \\ D_x f_2 & D_y f_2 \\ \vdots & \vdots \\ D_x f_M & D_y f_M \end{pmatrix} \begin{pmatrix} k_x \\ k_y \end{pmatrix} = \mathbf{D}\mathbf{k} = \mathbf{0} \tag{3}$$

where $D_x f_l = \partial f(\mathbf{r}_l)/\partial x$ and $D_x f_l = \partial f(\mathbf{r}_l)/\partial y$ with $\mathbf{r}_l$ being a node of a grid having M nodes on the patch. The matrix $\mathbf{D}$ is the set of gradients on the grid nodes, as indicated on the left in Eq. (3). using the continuous 2D Gaussian

$$g_{\sigma^2}(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{4}$$

the elements of $\mathbf{D}$, such as $D_x f_l$ and $D_y f_l$, can be preferably obtained by convolving the original discrete image with the discretized partial derivatives of the Gaussian. The parameter controlling the amount of smoothing the thus obtained *derivation filter* can apply is fixed by a certain $\sigma = \sigma_d$ in x and y directions as standard deviation, to avoid nonisotropic artificial bias. However, asking for nil (infinitesimal translation) error at every $\mathbf{r}_l$ with a common $\mathbf{k}$ may not be possible to fulfill in practice because f may not be perfectly linearly symmetric. The next best thing one can do is to solve the problem in the total least squares error sense such that $||\mathbf{D}\mathbf{k}||^2$ is minimized for a direction $\mathbf{k}$. The solution is given by the least significant eigenvector of the structure tensor, $\mathbf{S} = \mathbf{D}^T\mathbf{D}$, which is easy to obtain analytically as discussed in the following section. Alternatively, one can apply SVD numerically to $\mathbf{D}$ yielding the same solution obtained by an eigenvalue analysis of $\mathbf{S}$. Before computing the direction, in practice one needs to incorporate a window function $\mu_l = \mu(\mathbf{r}_l)$ into the solution as well because the patch must be cut-out of a larger image. This can be conveniently done in the tensor-space (at the level of the outer-product of the gradients) and via a Gaussian, to obtain a mathematically tractable optimization [2, 3].

$$\begin{aligned}\mathbf{S} &= \mathbf{D}^T\mathbf{D} \\ &= \begin{pmatrix} \sum_l (D_x f_l)^2 \mu_l & \sum_l (D_x f_l)(D_y f_l)\mu_l \\ \sum_l (D_x f_l)(D_y f_l)\mu_l & \sum_l (D_y f_l)^2 \mu_l \end{pmatrix} \\ &= \sum_1 (\nabla f_l \nabla^T f_l)\mu_l\end{aligned} \tag{5}$$

$$\begin{aligned}&= \lambda_{\max}\mathbf{k}_{\max}\mathbf{k}_{\max}^T + \lambda_{\min}\mathbf{k}_{\min}\mathbf{k}_{\min}^T \\ &= (\lambda_{\max} - \lambda_{\min})\mathbf{k}_{\max}\mathbf{k}_{\max}^T + \lambda_{\min}\mathbf{I}\end{aligned} \tag{6}$$

Here μ_l is a discrete Gaussian with a certain σ_w that defines the extension of the local fingerprint patches, $\lambda_{\max}$, $\mathbf{k}_{\max}$ are the most significant eigenvalue of $\mathbf{S}$ and its corresponding eigenvector, delivering the largest error and the maximum variation direction, respectively. Similarly the $\lambda_{\min}$, $\mathbf{k}_{\min}$ yield the

corresponding quantities for the least error and the direction of least variation respectively. Notice that $\mathbf{k}_{\max}$ and $\mathbf{k}_{\min}$ are always orthogonal ($\mathbf{S}$ is symmetric), have unit lengths, and sum to identity tensorially, $\mathbf{k}_{\max}\mathbf{k}^T{}_{\max} + \mathbf{k}_{\min}\mathbf{k}^T{}_{\min} = \mathbf{I}$. Thus to represent the direction we could relate it to $\mathbf{k}_{\max}$, the normal of the ridges/valleys, as well as to $\mathbf{k}_{\min}$ because knowing one determines the other. The representation of the direction is made by the tensor $\mathbf{k}_{\max}\mathbf{k}^T{}_{\max}$ rather than $\mathbf{k}_{\max}$ because the tensor representation will map the two possible numerical representations of the normal $\mathbf{k}$ and $-\mathbf{k}$ to the same (tensor) quantity avoiding the ambiguity inherent to vectors as representations of axes/directions.

Complex Representation of the Structure Tensor

There is a mathematically equivalent but a more convenient way of representing the structure tensor, by use of complex gradients [2, 4],

$$I_{20} = \sum_l (D_x f_l + iD_y f_l)^2 \mu_l = (\lambda_{\max} - \lambda_{\min}) e^{i2\varphi_{\max}} \tag{7}$$

$$I_{11} = \sum_l |D_x f_l + iD_y f_l|^2 \mu_l = \lambda_{\max} + \lambda_{\min} \tag{8}$$

with $\varphi_{\max}$ being the direction angle of $\mathbf{k}_{\max}$ and $i = \sqrt{-1}$.

The first benefit of complex representation is that the direction of the eigenvector is delivered by averaging (summation) squares of complex gradients, Eq. (6), in the argument of I_{20}, though in *double-angle representation* [5], and both eigenvalues are easily obtained by computing, $|I_{20}|$ and I_{11}. However easy to obtain, eigenvalues will not be necessary for many applications, as it is more useful to work with the sums and differences of them. This is because if $\lambda_{\min}$ is very small, an acceptable way to conclude upon this fact is to compare it with $\lambda_{\max}$. Accordingly, when we obtain a large (magnitude) complex number I_{20} for a patch, it means that we have a good direction fit (linearly symmetric patch) and a reliable estimate of the common direction will be found right in the argument of I_{20} (in ▶ double angle representation), with the reservation that $|I_{20}|$ must be close to I_{11}. By contrast, if the error of the worst direction is not much worse than the best direction then the direction fit is poor, making the corresponding argument angle meaningless automatically. Notice that $|I_{20}| \leq I_{11}$ and equality holds between the two quantities if and only if the iso-curve directions are aligned (linearly symmetric patch).

The next benefit is that the complex representation allows effective scale-space operations, including computation by subsampling, band-pass pyramids, extracting specific ridge frequencies (by changing σ_d, and σ_w), and coarse-to-fine refinements, etc. by using the complex image $(D_x f_l + D_y f_l)^2$ and its (realvalue) magnitude image, $|D_x f_l + D_y f_l|^2$.

Direction Fields as Features

The fact that scalar products on complex number fields is well defined makes direction fields *descriptive features* which can be used as complements to other descriptive features. If two fingerprints are registered, meaning that the query image f^q and the reference image f^r are rotated and translated such that they are aligned, then the scalar product between the corresponding direction fields of the query, $I_{20}(f^q)$, and the reference, $I_{20}(f^r)$, fingerprints

$$b(f^r, f^q) = \frac{|<I_{20}(f^r), I_{20}(f^q)>|}{\sqrt{<I_{20}(f^r), I_{20}(f^r)><I_{20}(f^q), I_{20}(f^q)>}} \tag{9}$$

can be used as a belief in the match. Here the scalar product is

$$<I_{20}(f^r), I_{20}(f^q)> = \sum_l I^*_{20}(f^r_l) I_{20}(f^q_l) \tag{10}$$

and the summation is applied either to a region, possibly weighted by some quality index [6–8], e.g. the common region of the fingerprint pair to be matched. The star as superscript denotes complex conjugation.

Direction Decomposition

A concept, i.e., closely related to direction fields is the decomposition of the original fingerprint in a set of images representing the (local) energy in quantized directions (typically 6–8 angles) and scales (typically 1–3 frequencies). Such decompositions can be obtained by a suitable Gabor filter bank independent of the direction field computations discussed earlier. Although the Gabor filter-bank filtered images can be interpolated to

generate accurate and dense direction fields [9], these have been mainly used to enhance fingerprints, and to estimate texture properties of fingerprints. The method suggested by [10] assumes that a landmark in each fingerprint of a pair to be matched is available or the pair is somehow registered with the corresponding landmarks. In regular concentric sectors of a circle (defined by a uniform polar grid of 5 radii and 16 angles) around the landmark, the average absolute deviations of Gabor-cosine filter responses (single frequency, eight directions) over the patch are computed. Called Fingercode, this set of texture measures constitutes a 640 dimensional ($5 \times 16 \times 8$) integer valued feature vector that can be used as a descriptive vector on its own or in conjunction with other features, Fig. 4.

Segmentation

In addition to their auxiliary or direct use to define descriptive features, the direction fields are also used in *segmenting fingerprints*. The latter refers to separating the image area that contains an acceptable quality of fingerprints, from the rest, typically the background. Because the fingerprint regions have a dominant orientation, meaning that there is a direction along which the gray-values change singificantly faster than the orthogonal direction, the absolute and/or relative differences of the structure tensor eigenvalues, λ_{min}, λ_{max} have been used to achieve segmentation [6, 11].

Minutiae

Minutiae are end-points of ridges or valleys of a fingerprint, occupying typically 0.1–0.5 mm on the skin, and are visible as 2–10 pixels in images captured at 500 dpi resolution. Minutiae are the most widely used features to match two fingerprints, for a variety of reasons, including that there is a great amount of human expertise in their use, and that it is difficult to reconstruct the original fingerprints only by the knowledge of minutiae, mitigating privacy concerns. A minutia can be of the type *termination* or *bifurcation*. A bifurcation of a ridge exists in conjunction with termination of a valley and vice-versa because the former engulfes the latter, by definition. This is known as *duality*. However, one must bear in mind that ridges appear as valleys and vice-versa depending on the sensing conditions, i.e., whether the dark pixels or the white pixels are ridges. Accordingly, the minutia-type, i.e., bifurcation or termination, as a descriptive feature is meaningful only if the interpretation ambiguity caused by sensing can be accounted for. Because from these two types of minutiae it is possible to derive other constellations, e.g., lake, spur, crossover, Fig. 2, several national agencies relying on

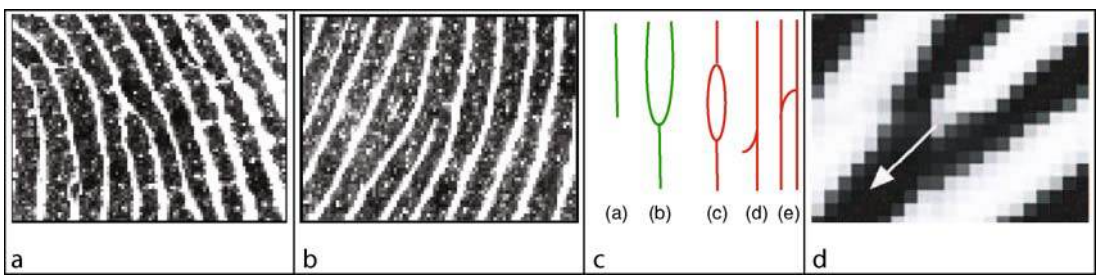
Fingerprint Features. **Figure 1** Commonly used classes to categorize fingerprints [27]. (**a**) Arch, (**b**) Tented Arch, (**c**) Left Loop, (**d**) Right Loop, (**e**) Whorl, (**f**) Twin Loop.

minutiae for their fingerprint processing base their taxonomy only on termination and bifurcation, e.g., FBI in USA [12]. Before minutiae extraction, *fingerprint enhancement* is applied if fingerprints are deemed noisy, usually according to an automatically extracted quality measure [7, 8, 13, 14].

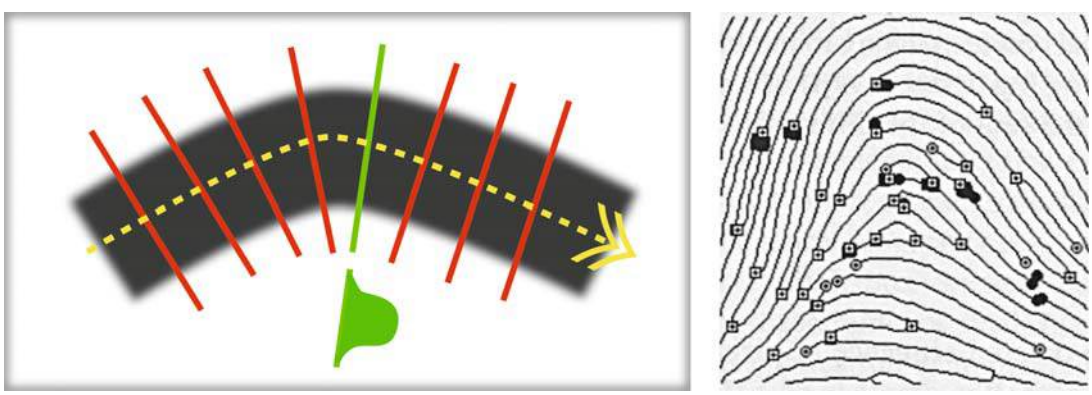
Two main ways of minutiae extraction can be achieved by (1) by binary image processing operations, (2) by using gray-value image processing techniques.

Assuming that the binary image of a fingerprint can be obtained and it has a reasonably high fidelity w.r.t. ridges, ► fingerprint thinning can be achieved by morphological operators (errosion and dilation) or by distance transforms [15–17]. A number of algorithms to extract minutiae from skeletonized binary images exist. It is common that at the beginning, there are several thousands of minutiae candidates of which only approximately 50 are real. Various criteria for validating the endpoints, including the duality, a minimum length of the ridge or valley , are used to suppress spurious false minutiae [18].

However, minutiae detection based on binary images has a shortcoming, lack of robustness when used for low quality fingerprint images. Because ridge skeletons are obtained by applying a thinning method to the binarized fingerprint, the binary ridges should correspond to real ridges accurately if thinning procedure is to be successful. This puts high demands on the quality of the fingerprints, as well as the adaptiveness of the binarization since the resulting binary ridges might not represent the real ridges sufficiently well. Extracting minutiae from gray images, without passing through binarization, offers better opportunities in this respect. The ridges can be directly followed in the gray-value image by use of the direction field, and the gray-value ridge profiles [6, 11], Fig. 3.



Fingerprint Features. **Figure 2** Illustration of minutiae types and duality. (**a**) a ridge termination engulfed in a valley bifurcation; (**b**) vice-versa. (**c**) basic ridge types in green (termination, bifurcation) and derived types in red (lake, spur, crossover) (**d**) the direction of a minutia exemplified at a ridge-bifurcation.



Fingerprint Features. **Figure 3** Illustration of thinning and minutia detection by ridge following in gray-images [11]. On the *left*, a segment of a ridge is represented. Gray-value profiles, like the one in green, are regularly sampled and tracked along the ridge, until a termination or a bifurcation is found. On the *right*, the result is shown where the white circles and squares represent terminations, and bifurcations respectively. The black circles and boxes are improvements of a postprocessing.

Alternatively, a large number of candidate minutiae can first be obtained, e.g., by detecting lack of linear symmetries during the direction field estimation, then a gray-value model of the minutiae, e.g., the parabolic appearance of terminations and bifurcations, can be enforced the candidates to retain the true minutiae [19], Fig. 4.

Minutia Direction When matching or registering two fingerprints the ▶ minutia direction is a valuable discriminative information. The minutiae directions can be either extracted from the direction field directly or from the direction of the binarized and thinned ridges, corresponding to minutiae locations, Fig. 2. The directions along with the type information (termination or bifurcation) are attached to minutiae coordinates.

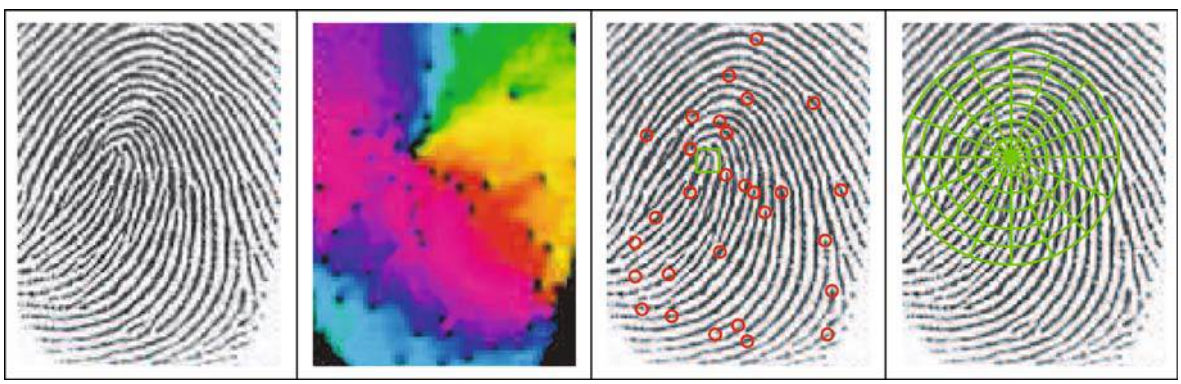
Spatial Frequency Another descriptive feature which can be attached to minutiae positions is the spatial frequency information in the vicinity of minutiae. The spatial frequency is usually defined in terms of a direction in fingerprints and has different implementations [20]. One implementation is to use the average frequency of the ridge or the count of ridges in a fixed line segment orthogonal to the minutia direction. Another implementation of the frequency measure is to count ridges or the average frequency along the line joining a pair of minutiae. Because pairs as well as triplet constallations of minutiae are commonly used in fingerprint matching, the frequency measures are attached as a descriptive feature to the corresponding, pairs or triplets.

Singular Points

Singular points are landmarks that are defined in large image patches (1–5 mm) compared to the size of minutiae. There are typically 1–2 singular points in a fingerprint though they may occasionally be missing or may be difficult to identify in a fingerprint. Three basic types can be discerned, loop (also known as core), whorl, and delta.

A major use of them is to classify a fingerprint typically into one of the six categories, (Left-loop, Right-Loop, Double-Loop, Arch, Tented-Arch, Whorl) which are different constellations of loops and deltas, Fig. 1. Such rough categorizations are employed to match, and to organize massive amounts of fingerprits data efficiently.

Loops can provide a unique intrinsic global orientation and position for a fingerprint, allowing an orientation and translation normalization of the fingerprint only on the basis of itself. Most whorls and deltas can provide a direction too, though these are in general not unique. Two singular points in the same image provide always a unique ▶ intrinsic direction of fingerprint. This normalization is a practical alternative to registration by minutiae or can be complementary. Every



Fingerprint Features. **Figure 4** Illustration of a use of direction field. From the *left*, the original image, the direction vector field color coded, the original superimposed with minutiae locations and the loop singularity [19, 24] are shown, respectively. On the far *right* the Fingercode grid placed, on the loop singularity, is shown [10]. The direction field image (*second*) represents the complex quantities I_{20} (7), where the argument of I_{20} is mapped to the Hue (*color*) of the HSV color model (same color indicates common direction) and the magnitude representing the quality of the direction fit is mapped to the Value (intensity).

fingerprint (the query, as well as every fingerprint in the database) is rotated and translated such that a reference point and a half-line that is well defined w.r.t. a singular point of the fingerprint become the origin and the positive x-axis. Two translation and rotation normalized fingerprints are then more efficiently matched – with minutiae or other features, because no rotation or translation compensation specific to the considered pair will be necessary.

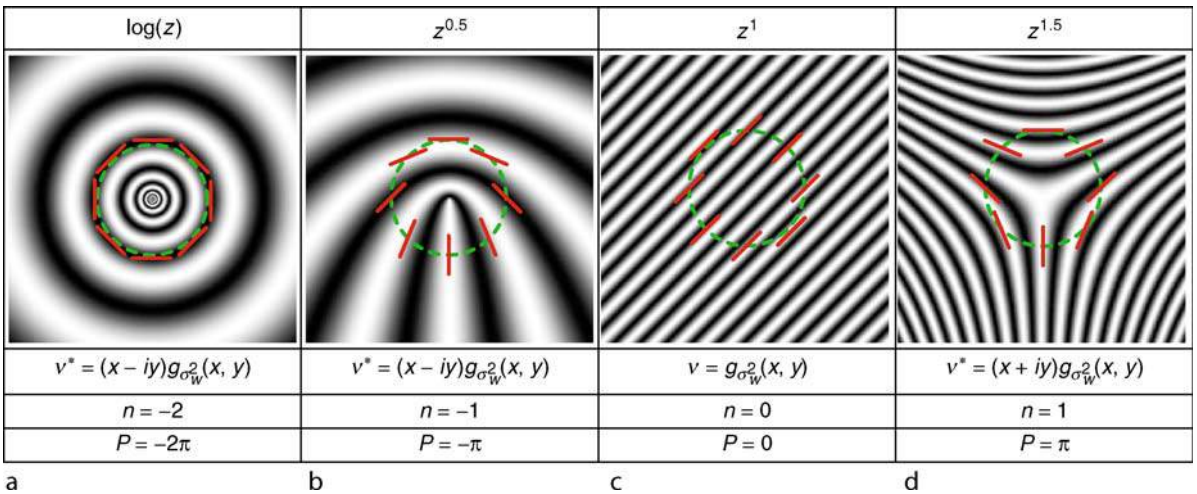
Finally, singular points can function as anchors to extract other descriptive features, e.g., the spatial frequency. One can count the ridges along a line joining two singular points, or along a line joining a minutia and a singular point, etc. The spatial frequency estimation issue is analogous to the one that has been discussed in conjunction with minutiae.

Singularities by Poincaré Index

One of the oldest singular point detection techniques used in fingerprint processing is the Poincaré index [21]. The index is defined for a path in a vector field and represents the total amount of angle change of the vectors along the curve. Assuming that the curve is closed and it is in the gradient field of a fingerprint then the Poincaré index, P, is given by

$$P = \oint \frac{\partial \theta}{\partial x} dx + \frac{\partial \theta}{\partial y} dy = \int \int \left(\frac{\partial^2 \theta}{\partial x \partial y} - \frac{\partial^2 \theta}{\partial y \partial x} \right) dx dy \tag{11}$$

where the function $\theta(x, y)$ represents the argument (angle) of the gradient vectors and the last expression is obtained by Green's Lemma. It is worth noting that even though the original fingerprint image is assumed differentiable (continuous) the gradient angle is not continuous, (π and $-\pi$) though its partial derivatives are. By laying the closed curve around a loop, a whorl, a regular (non-singular) point, and a delta, it can be concluded that P will assume $-2\pi, -\pi$, 0 and π radians, respectively. In Fig. 5 stylistic models of such fingerprint patches are shown along with segments of iso-curves (to which the gradients are orthogonal but are not shown for convenience). When one walks the dashed circle in full, the direction of iso-curves, and thereby the gradient angles change with the Poincaré index. This observation is used, typically along with the curve integral of Eq. (11), to determine if a candidate point is a whorl, loop, regular, or delta type. It is also possible to compute P according to the right hand side of the equation, by a double integral applied to the interior patch of the curve. By using the directions of linear-symmetry vector field, as opposed to those of the gradient field and the double integral [22]



Fingerprint Features. Figure 5 The *top row* shows the harmonic functions that generate the iso-curves of the patterns in the *second row*. The iso-curves (their linearized examples are shown as red line segments) are given by a weighted combination of the real and the imaginary parts of the respective harmonic functions, with a certain ratio between the weights, defining the direction parameter φ of each pattern [23]. The *third row* shows the filters, v where $g_{\sigma_w^2}$ is an ordinary Gaussian (4), to detect the singularity points and φ by a (complex) convolution applied to the direction field (12). The *third row* shows the symmetry order of the filters. The last row shows the Poincaré index of gradients.

suggested an alternative way of computing P. In this case the angles of the used vector field are continuous from the beginning so that no special care needs to be taken to achieve continuity at angles around π and $-\pi$. The resulting P must be divided by 2 to correspond to the gradient based Poincaré index.

Singularities by the Generalized Structure Tensor

A singular point can also be detected by use of the Generalized Structure Tensor (GST), which is an extension of the structure tensor to curvilinear (harmonic) coordinates [9, 23]. The fundamental idea is the same as that of the structure tensor – to find an (unknown) angle such that the patch remains invariant to a small translation along the found angle direction but in the curvilinear coordinates. It turns out that in this model, a singularity can be detected by complex filtering of the direction fields, already in the complex representation (7).

$$I'_{20} = \sum_l (D_x f_l + iD_y f_l)^2 v_l^* = (\lambda_{\max} - \lambda_{\max}) e^{i2\varphi_{\max}} \quad (12)$$

Here v_l is a filter specialized to detect a loop, a delta or a whorl, Fig. 5. The magnitude of a filter response, which is complex valued, encodes the likelihood that a location represents a singularity exactly in the same way as the ordinary structure tensor, but now the coordinates are harmonic, representing a pattern of a singularity, and the λ_{max}, and λ_{min} are the error extrema due to translation in curvilinear trajectories having a certain direction. Likewise, its argument (angle) encodes the intrinsic orientation of the singularity (for loops and deltas their global inclination, for whorls the amount of chirality). The singularity filters can be implemented by derivatives of Gaussians which are separable, making them 1D filters. Because the complex feature space obtained from such filter responses are continuous both in their arguments and positions, scale-space filtering, e.g., coarse-to-fine refinement, is possible [24]. That the symmetry axes (intrinsic orientation) are available in the GST method is useful, because the obtained angle information can be used as a descriptive feature attached to the singular point coordinates, much like the use of minutiae orientations in fingerprint matching. Additionally, loop orientations alone allow a normalization/registration of a fingerprint pair even if other singular points lack, and no minutiae are available.

Singularities by Other Methods

The methods discussed earlier can find singular points by modeling direction variations on closed curves (in practice a circle) or in regions containing a singularity. Methods which do not use closed paths are exemplified as follows. Such a method to obtain singularities is the early suggestion by [25] which models the direction variations along the horizontal scan lines. Information defining the location and the type of the singularity is contained in the direction information around the singular point and the horizontal lines contain only a part of this. This information is instead injected into the model in terms of orientation-change rules between scan lines. In [26], gradient vectors model half a circle, like "n." Then generalized Hough transform is used to find a peak, suggesting the location of a loop. In contrast to GST and Poincaré index methods, the (loop) inclination is assumed to be approximately vertical, or a separate model is designed for alternative loop inclinations.

Summary

Descriptive features are used to match fingerprints. They include the locations of minutiae points, and the singular points. The location information can be enhanced with additional descriptive measurements including the local direction of the ridges and valleys at minutiae locations, the intrinsic orientation of singular points, the type of the singular points, ridge counts or average frequencies between minutiae as well as singular points. To extract such descriptive information direction maps are computed. Being texture measures, structure tensor representations of direction maps can also be used as descriptive features on their own if anchor points are available or in addition to minutiae based features. Similarly, Gabor filters can be used to obtain descriptive features if anchor points are available. Commonly used anchors for registration as well as descriptive features are the three basic singularity types, loops, whorls, and deltas. They can be detected and described independent of minutiae information.

Related Entries

- ▶ Fingerprint Enhancement
- ▶ Fingerprint Matching
- ▶ Fingerprint Quality
- ▶ Fingerprint Registration
- ▶ Fingerprint Sensing

References

1. Locard, A.: L'Identification des Récidivistes. A. Maloine, Paris (1909)
2. Bigun, J., Granlund, G.: Optimal orientation detection of linear symmetry. In: First International Conference on Computer Vision, ICCV, London, June 8–11, pp. 433–438. IEEE Computer Society, London (1987)
3. Kass, M., Witkin, A.: Analyzing oriented patterns. Comput. Vision Graph. Image Process. **37**, 362–385 (1987)
4. Bigun, J., Granlund, G., Wiklund, J.: Multidimensional orientation estimation with applications to texture analysis and optical flow. IEEE-PAMI **13**(8), 775–790 (1991)
5. Granlund, G.: In search of a general picture processing operator. Comput. Graph. Image Process **8**(2), 155–173 (1978)
6. Ratha, N.K., Chen, S., Jain, A.K.: Adaptive flow orientation-based feature extraction in fingerprint images. Pattern Recogn. **28**(11), 1657–1672 (1995). URL http://dx.doi.org/10.1016/0031-3203(95)00039-3
7. Grother, P., Tabassi, E.: Performance of biometric quality measures. IEEE Trans. Pattern Anal. Mach. Intell. **29**(4), 531–543 (2007). URL http://dx.doi.org/10.1109/TPAMI.2007.1019
8. Fronthaler, H., Kollreider, K., Bigun, J., Fierrez, J., Alonso-Fernandez, F., Ortega-Garcia, J.: Fingerprint image quality estimation and its application to multi-algorithm verification. IEEE Trans. Inform. Forens. Security **3**(2): 331–338 (2008)
9. Bigun, J.: Vision with Direction. Springer, Heidelberg (2006)
10. Jain, A.K., Prabhakar, S., Hong, L., Pankanti, S.: Filterbank-based fingerprint matching. IEEE Trans. Image Process. **9**(5), 846–859 (2000). URL http://dx.doi.org/10.1109/83.841531
11. Maio, D., Maltoni, D.: Direct gray-scale minutiae detection in fingerprints. IEEE Trans. Pattern Anal. Mach. Intell. **19**(1), 27–40 (1997). URL http://www.computer.org/tpami/tp1997/i0027abs.htm
12. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, Berlin (2003). URL http://bias.csr.unibo.it/maltoni/handbook/
13. Hong, L., Wand, Y., Jain, A.: Fingerprint image enhancement: Algorithm and performance evaluation. IEEE-PAMI **20**(8), 777–789 (1998)
14. Chen, Y., Dass, S.C., Jain, A.K.: Fingerprint quality indices for predicting authentication performance. In: Audio- and Video-Based Biometric Person Authentication, p. 160 (2005). URL http://dx.doi.org/10.1007/11527923_17
15. Xiao, Q., Raafat, H.: Fingerprint image postprocessing: A combined statistical and structural approach. Pattern Recogn. **24** (10), 985–992 (1991). URL http://dx.doi.org/10.1016/0031-3203(91)90095-M
16. Hung, D.C.D.: Enhancement and feature purification of fingerprint images. Pattern Recogn. **26**(11), 1661–1671 (1993). URL http://dx.doi.org/10.1016/0031-3203(93)90021-N
17. Shih, F.Y., Pu, C.C.: A skeletonization algorithm by maxima tracking on Euclidean distance transform. Pattern Recogn. **28** (3), 331–341 (1995)
18. Farina, A., Kovacs Vajna, Z.M., Leone, A.: Fingerprint minutiae extraction from skeletonized binary images. Pattern Recognition **32**(5), 877–889 (1999). URL http://www.sciencedirect.com/science/article/B6V14-3WMK59F-D/2/bf21218ba618c9f63efb1663ea24a6f6
19. Fronthaler, H., Kollreider, K., Bigun, J.: Local feature extraction in fingerprints by complex filtering. In: S.Z.Li et al. (ed.) International Workshop on Biometric Recognition Systems – IWBRS 2005, Beijing, Oct. 22–23, LNCS 3781, pp. 77–84. Springer, Heidelberg (2005)
20. Maio, D., Maltoni, D.: Ridge-line density estimation in digital images. In: International Conference on Pattern Recognition, vol I, pp. 534–538 (1998). URL http://dx.doi.org/10.1109/ICPR.1998.711198
21. Kawagoe, M., Tojo, A.: Fingerprint pattern classification. Pattern Recogn **17**, 295–303 (1984)
22. Bazen, A., Gerez, S.: Systematic methods for the computation of the directional fields and singular points of fingerprints. IEEE-PAMI **24** (7), 905–919 (2002)
23. Bigun, J., Bigun, T., Nilsson, K.: Recognition by symmetry derivatives and the generalized structure tensor. IEEE-PAMI **26**, 1590–1605 (2004)
24. Nilsson, K., Bigun, J.: Localization of corresponding points in fingerprints by complex filtering. Pattern Recogn. Lett. **24**, 2135–2144 (2003)
25. Wegstein, J.H.: An automated fingerprint identification system. Tech. Rep. Special Publication 500-89, National Bureau of Standards, NBS (1982). URL http://www.itl.nist.gov/iad/894.03/fing/Special_Publication_500-89.pdf
26. Novikov, S., Kot, V.: Singular feature detection and classification of fingerprints using Hough transform. In: E. Wenger, L. Dimitrov (eds.) Proc. of SPIE, vol. 3346, pp. 259–269 (1998)
27. Garcia, J.O., Aguilar, J.F., Simon, D., Gonzalez, J., Zanuy, M.F., Espinosa, V., Satue, A., Hernaez, I., Igarza, J.J., Vivaracho, C., Escudero, D., Moro, Q.I.: MCYT baseline corpus: a bimodal biometric database. IEE Proc. Vision Image Signal Process. **150** (6), 395–401 (2003). URL http://ieeexplore.ieee.org:80/xpls/abs_all.jsp?isNumber=2825%2&prod=JNL&arnumber=1263277&arSt=+395&ared=+401&arNumber=1263277

Fingerprint Identification

- ▶ Fingerprint Indexing
- ▶ Fingerprint Recognition, Overview

Fingerprint Image Compression

► Fingerprint Compression

Fingerprint Image Enhancement

Masanori Hara
NEC Corporation, Tokyo, Japan

Synonyms

Fingerprint contrast enhancement; Ridge enhancement; Ridge extraction

Definition

Fingerprint image enhancement is the process of applying techniques to emphasize fingerprint images in order to facilitate the identification of ridge valley structures and hence their features.

Introduction

Computerized fingerprint feature extractors more or less require some sort of image pre-processing or enhancement to improve perceptibility. In doing so, they need to contend with two major types of problems: one is associated with image contrast such as insufficient dynamic range, and the other is associated with adverse physical factors such as scars, blurs, creases, sweat pores, and incipient ridges. Fingerprint image enhancement aims to minimize the undesired effects caused by such elements in order to extract a sufficient number of reliable features, namely, minutiae and ► fingerprint singularities (cores and deltas). Broadly speaking, fingerprint image enhancement encompasses, but is not limited to, the intermediate steps such as contrast enhancement, pore and incipient ridge removal, ridge orientation and frequency estimation, foreground segmentation, and ridge enhancement filtering.

The focus here is on the performance and limitations of current image enhancement techniques rather than on their algorithmic details. For this purpose, many samples including problematic images and their corresponding enhanced images are presented.

Fingerprint Image Digitalization and Density

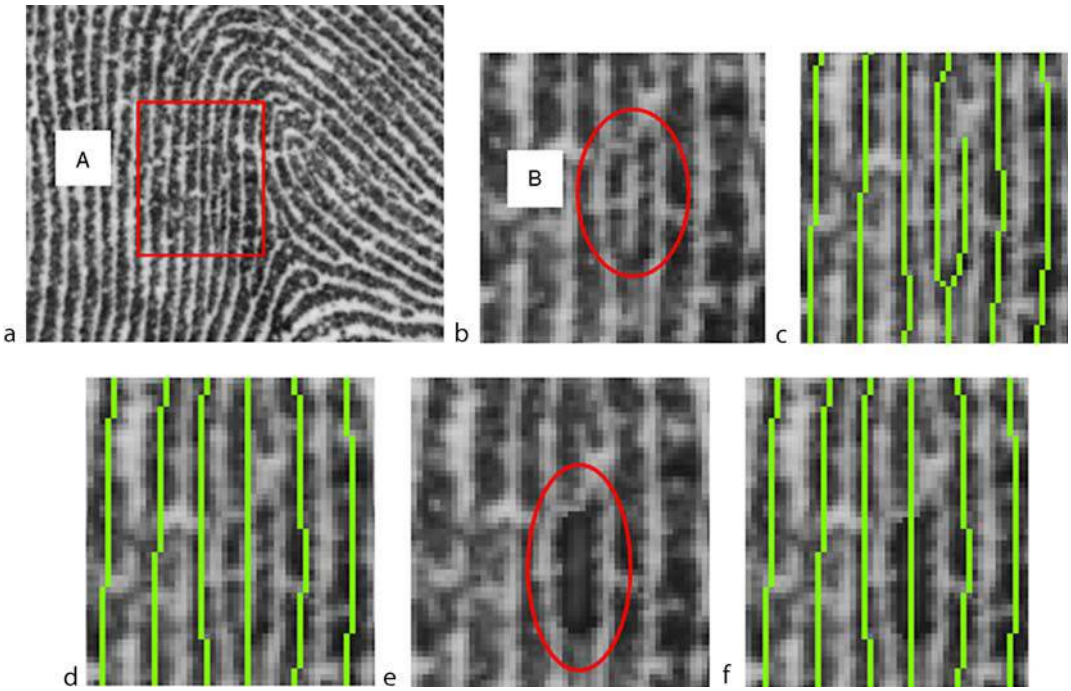
Fingerprint images are digitized through either inked-print scanning or live scanning, most often with a resolution of 500 dpi and a depth of 8 bits (i.e., 256 gray levels) in compliance with the NIST standard [1]. The gray level is a value associated with each pixel representing its intensity or luminance. However, the term *density*, the degree of ink thickness on the paper surface, has been used throughout this section for the sake of illustration. Thus, the higher the density, the darker the ridges, and vice versa.

Ideally, the density of pores should be higher than that of valleys and the density of incipient ridges should be lower than that of true ridges. In fact, this is precisely what some feature extractors traditionally use to distinguish pores and incipient ridges from true ridges and valleys. Although most inked-print scanned images have continuous density, some live scanned images exhibit a sparse nature. It has been reported that the effective bit depth of some live scanners is only 2 or 3 bits [2, 3]. Obviously, such loss of information makes the subsequent processes virtually impossible to distinguish the key features.

Recognition of Fingerprint Ridge

Since ridge orientation and frequency characterize the local ridge valley structure in the region of interest, the problem of fingerprint ridge recognition essentially simplifies to the task of estimating these two pieces of information. Therefore, local ridge orientation and frequency estimation play a key role in fingerprint ridge recognition.

Local orientation estimation, taking advantage of the fact that ridge orientation does not change suddenly when viewed locally, can "interpolate" ridge orientation even in obscured regions. However, this is not the case in frequency estimation; frequency can become rather unstable when there is a sudden change



Fingerprint Image Enhancement. Figure 1 An Example of Unstable Frequency Estimation (NIST DB#27 002T). Notes: (**c**) and (**f**) A skeleton image was extracted by one of the traditional algorithms [7] using the contextual ridge enhancement filter with a narrow spacing, which corresponds to higher frequencies. (**d**) A skeleton image was extracted by the same algorithm as in (**c**) and (**f**) with a wide spacing, which corresponds to lower frequencies.

in ridge spacing even in a clear, well-defined region, as illustrated in Fig. 1. The ridges in the region marked by the red square (region A) in Fig. 1(a) are nicely aligned in the vertical direction, whereas frequency significantly changes due to the presence of a spur, alternatively called a whisker, in the region marked by the red oval (region B) in Fig. 1(b). An average inter-ridge spacing is 8.7 pixels (0.435 mm) in region A whereas it is only 5.7 pixels (0.285 mm) in region B, which is narrower than the neighboring region by 35%. For this image, some feature extractions are able to correctly extract this narrow spur as shown in Fig. 1(c) but some others fail (Fig. 1(d)).

Even if the spur in region B is invisible as shown in Fig. 1(e), it is still easy to estimate ridge orientation in the region with a high degree of certainty. However, if this small region is contaminated with noise, most feature extractions incorrectly estimate the frequency in region B to be the same as that of its neighbor, which results in a failure to detect the spur (Fig. 1(f)). An ideal feature extractor should be able to mark this region as "indeterminate" because it is difficult even for human examiners to identify the spur confidently.

Intermediate Steps in Fingerprint Image Enhancement

A typical set of intermediate steps in fingerprint image enhancement includes:

1. Contrast enhancement or normalization.
2. Pore and incipient ridge removal.
3. Ridge orientation estimation.
4. Frequency estimation.
5. Foreground segmentation.
6. Ridge enhancement filtering.

Contrast Enhancement

Whatever features or structures there may be, either local or global, distinctiveness is important to appropriately separate one from the other. The conditions that are preferably satisfied may include uniform background density and a sufficiently wide dynamic range between ridges and valleys/background. If these conditions are fulfilled, a simple stretching and/or

thresholding should suffice. In reality, however, more elaborate and rigorous approaches are needed. Some real issues related to dynamic range are outlined in the following:

1. Uneven dynamic range.
   A sample of uneven dynamic range is presented in Fig. 2(a). The ridges on the left (surrounded by the red oval) are substantially lighter than the ones on the right.
2. Uneven valley density.
   A sample of uneven valley density is shown in Fig. 2(b). This is a latent image that is lifted from paper. The latent print is impressed on across the regions where letters ("O", "A", "N") are printed. Here, the ridges and valleys cannot be recognized easily because their actual density considerably deviates from their original definition in which "the ridges are dark and the valleys are light"; the density of the valley on the letters is contrarily high and local dynamic range is extremely narrow, whereas the valley density in the plain region is low.
3. Noisy background.
   The background containing leftover fingerprint images or stripe patterns resembling fingerprints makes it difficult to isolate true fingerprint patterns. Problematic live scanned images and an inked image are presented in Fig. 2(c, e) and 2(d), respectively.

Contrast enhancement is a technique to accommodate such problems by expanding the dynamic range of ridges and valleys. Adaptive histogram equalization is a popular contrast enhancement technique. Other popular techniques include a simple linear contrast stretching that uses the local minimum and maximum densities, and a density normalization that uses the local density mean and variance [2, 5]. Contrast-enhanced images of the sample problematic images are presented in Fig. 2 (a′) through (e′).

Although contrast enhancement is a powerful tool, it has a drawback, i.e., it boosts background noise at the same time since it cannot selectively enhance only the targeted region unless some additional information is given. As shown in Fig. 2(a′), (c′), (d′), and (e′), the distinguishability of the foreground and background is lower than the original.

However, it should be stressed that it is still imperative to employ contrast enhancement when dealing with poor quality images, mainly latent images (Fig. 2(b′), for example). Once the ridge valley structure becomes visible, it essentially boils down to the problem of identifying and analyzing ridge continuity.

Pore and Incipient Ridge Removal

Sweat pores are major obstacles in frequency estimation. There are a variety of methods for removing pores or at least for reducing their side effects. Some methods do not remove pores but they remove false minutiae that possibly originated from the pores. Other methods rely on the fact that pores are enclosed by darker pixels as shown in Fig. 3(a), and they can remove typical pores but not problematic pores such as continuous pores and swollen pores as shown in Fig. 3(b). Ridge structures such as lakes and spurs are easily confused with pores, leading to miscalculation of frequency if they are falsely filled in.

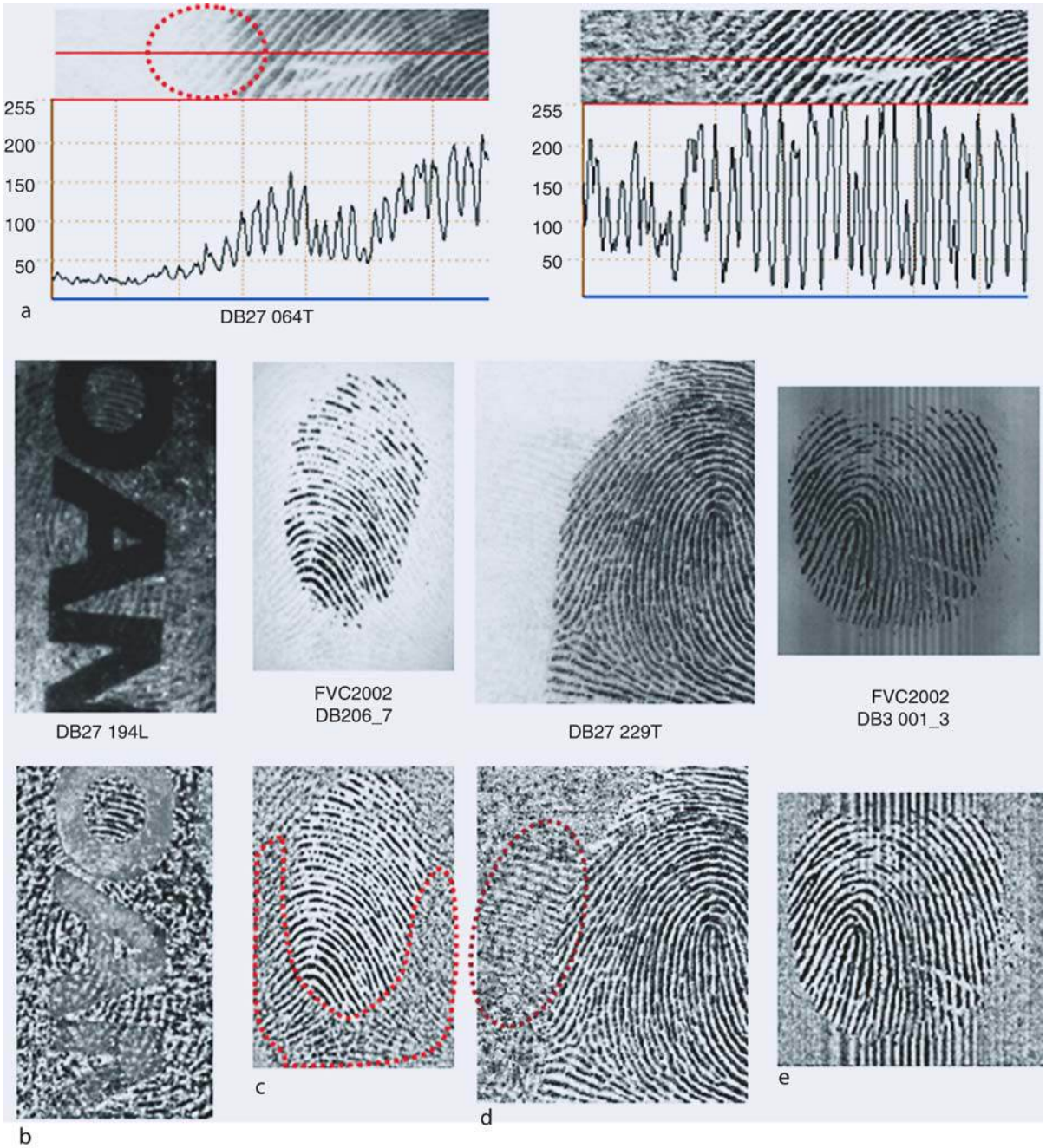
Incipient ridges are another obstacle in frequency estimation. The significant incipient ridges as shown in Fig. 3(c) can easily fool frequency estimation algorithms.

These two factors have not yet been fully explored, and their distinguishability plays an important role in improving fingerprint matching accuracy.

Ridge Orientation Estimation

Ridge orientation estimation is a fingerprint-specific image processing technique. A ridge orientation estimation algorithm was developed for a FBI system in the 1960s. In the 1960s and 1970s, many ridge orientation estimation algorithms set "slits" of predetermined orientations (8, 12, or 16 quantized orientations) and analyzed the density response [6–8]. The orientation slit having a higher amount of density change is indicative of the slit running perpendicular to the direction of the ridge flow. Similarly, the slit with a lower amount of density change is indicative of the slit running parallel to the direction of the ridge flow.

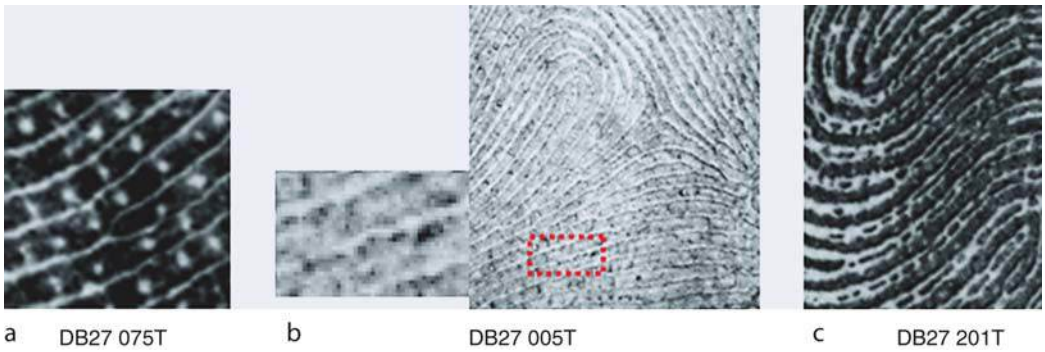
In the 1980s, more sophisticated methods were introduced to extract ridge orientation such as a method based on the gradient of two-dimensional vectors whose components are derivatives of densities at horizontal and vertical orientations [2, 5], and a method based on the two-dimensional Fourier transform [9, 10].



Fingerprint Image Enhancement. Figure 2 Examples of Contrast-related Problematic Images. Notes: (**a'**) through (**e'**) The images were contrast-enhanced with one of the local adaptive stretching methods specialized for fingerprints [4].

In this process, the confidence level of ridge orientation is calculated. The difference in density fluctuation between the estimated orientation and its orthogonal orientation can be a base for confidence, and the power spectrum is another in the case of the Fourier analysis.

Since all these techniques estimate ridge orientation locally, the influence of adverse factors such as scars and smudges are not negligible and often lead to wrong estimation. In order to correct such anomaly, local orientation is examined for validity and re-estimated from its neighbor. This process is called



Fingerprint Image Enhancement. **Figure 3** Pores and Incipient Ridges.

ridge orientation smoothing, and several such techniques have been proposed [2, 5, 9, 11].

Ridge orientation smoothing also has a drawback. The orientation in the region where the ridge flow is not stable (e.g., in the proximity of the core and delta) cannot be properly corrected due to its interpolative nature. It also fails and propagates errors if the overall estimation quality is low because it is based on the assumption that the majority of orientations of neighboring regions are indeed correct.

Examples of problematic images in orientation estimation are presented in Fig. 4. Orientations in the red oval in Fig. 4(a′) and (b′) are incorrectly estimated because of smudges and fragmented ridges.

One of suggested methods to improve estimation accuracy is to use global pattern types and prior knowledge of ridge flow. Once the core and delta have been extracted with high confidence, the global pattern shape can be estimated. This information can help estimate and adjust local ridge orientation more accurately.

Frequency Estimation

Frequency estimation is another fingerprint-specific image processing technique. Frequency is defined as the number of ridges per unit length and is often interchangeably referred to as the inverse of the inter-ridge distance. It is far more difficult to estimate than orientation, and that explains why most feature extractions in the 1960s and 1970s did not fully exploit this information.

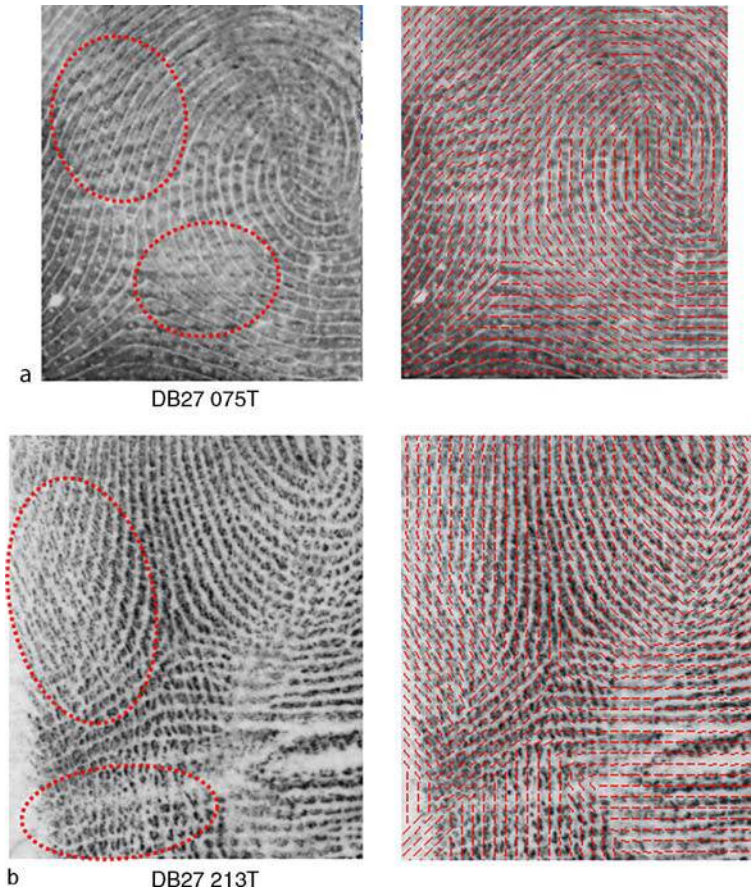
In the 1980s, frequency analysis such as the two-dimensional Fourier transform was proposed to estimate frequency [9, 10]. Another technique of frequency estimation analyzes peak intervals from gray-level profile orthogonal to the ridge orientation [2, 5].

In Fig. 5(a) an example of a problematic image with a sudden frequency change, denoted by the red oval is presented. In Fig. 5(b), the true frequency is reflected via some manual correction, and Fig. 5(c) shows an example of automatically estimated frequency image using one of the latest algorithms [12]. Dark density pixels correspond to the region where the inter-ridge spacing is narrow, that is, frequency is high. It can be observed that the frequency of the area with very narrow inter-ridge spacing is falsely estimated to be halved from its true value.

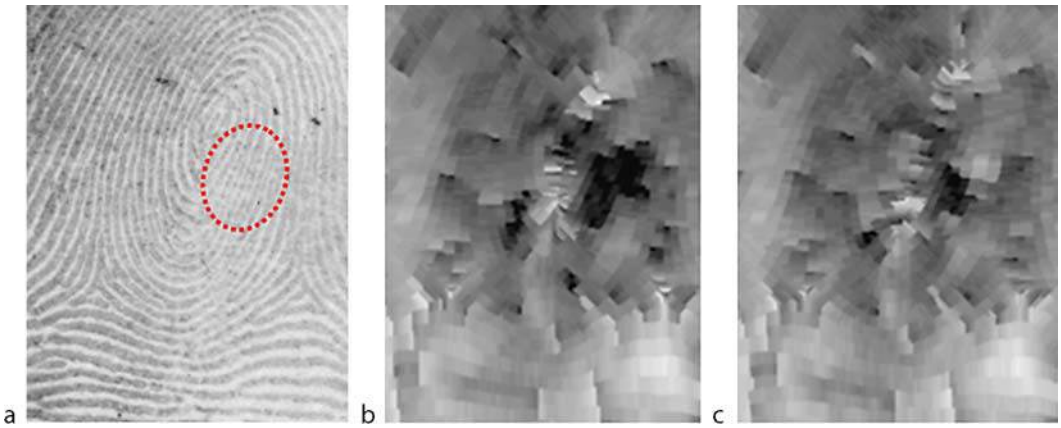
It is known that the presence of a minutia affects the structure of its surroundings and hence the corresponding local frequency. This often becomes a problem in frequency estimation where a strong frequency smoothing aimed to facilitate the estimation process can adversely eliminate true minutiae.

Foreground Segmentation

It is natural to conduct minutia extraction only in the foreground region to minimize the possibility of extracting false minutiae. Foreground segmentation aims to distinguish the fingerprint ridge region from the background. Some methods rely heavily on the confidence of the ridge orientation to define the foreground, whereas others rely on gray-level statistics as well. As already explained, gray-level analysis is not an ideal approach when dealing with very low quality images such as the ones shown in Fig. 2.



Fingerprint Image Enhancement. Figure 4 Problematic Images in Ridge Orientation Estimation. Notes: (**a′**) and (**b′**) A ridge orientation image was extracted by one of the traditional algorithms [7].



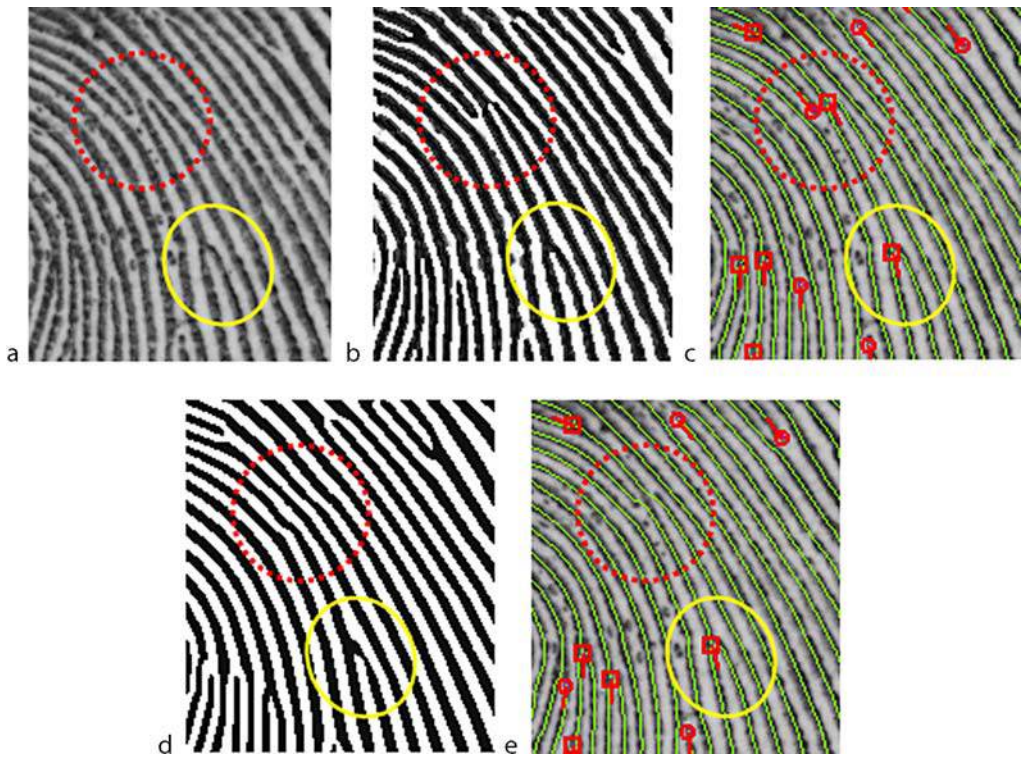
Fingerprint Image Enhancement. Figure 5 Problematic Images in Frequency Estimation (NIST DB#27 073T). Notes: (**b**) True frequencies were calculated from an ideal skeleton image, which was manually generated so that skeleton curves correctly coincided with the original ridges. (**c**) Frequencies were calculated from an automatically extracted skeleton image using one of the recent algorithms.

Ridge Enhancement Filtering

Ridge enhancement filtering is another fingerprint-specific image processing technique. In the 1960s and 1970s, most filtering techniques were labeled contextual. They used filtering masks similar to the ones used to estimate orientation, either fixed size or predetermined variable frequencies. However, it was difficult for these techniques to flexibly adapt to very narrow or very wide ridges and spacing [6–8]. In the 1980s, a more sophisticated method based on the two-dimensional Fourier transform was proposed [9, 10]. In the 1990s and 2000s, Gabor filtering and wavelet filtering were introduced [2, 5, 13].

Conceptually, ridge enhancement filtering aims to "enhance" ridges by generating stripe patterns from scratch using the previously estimated orientation and frequency. Strong enhancement is effective for low quality images but at the risk of destroying the original ridge structure. The strength of filtering thus needs to be controlled adaptively and depends on the field in which it is used: law enforcement and non-law enforcement. In the former case, the original ridge structure needs to be preserved as much as possible in order to improve compatibility with the examiners' definition of minutiae since it still relies on manual processing such as latent minutia coding. This is important to improve latent-print matching accuracy, especially for fragmental latent prints with few minutiae. In order to match such latent prints, even unstable minutiae need to be incorporated to increase chances of hit. On the contrary, in the latter case, which is fully automatic, neither the original ridge structure has to be preserved nor is compatibility with the examiners' definition critical.

With respect to minutia preserving ability, there are two types of minutiae to be considered: stable minutiae and unstable minutiae. The stable minutia is a minutia that is topologically isolated from other minutiae with no chance of interfering with other minutiae.



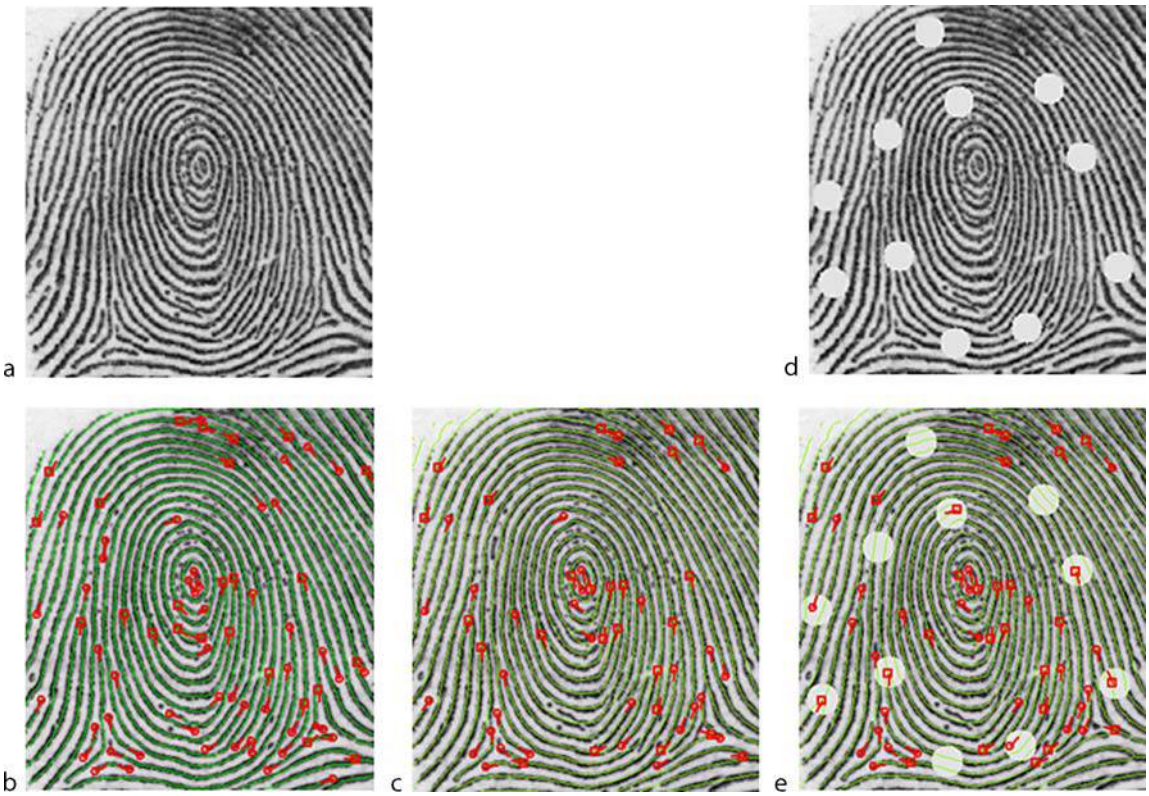
Fingerprint Image Enhancement. **Figure 6** Unstable and Stable Minutiae (NIST DB#27 076T). Notes: (**b**) The ridge image (**a**) was enhanced by one of the popular algorithms [9] with a relatively weak enhancement parameter. (**c**) A skeleton image and minutiae were automatically extracted from the image in (**b**). (**d**) The image in (**b**) was enhanced by one of the popular algorithms [9] with a relatively strong enhancement parameter. (**e**) A skeleton image and minutiae were automatically extracted from the image in (**d**).

The unstable minutia is a minutia that may either remain unchanged or completely disappear depending on the physical conditions of its surroundings. Crossover minutiae are a typical example of unstable minutiae. In Fig. 6(a) a stable minutia is presented in the yellow circle and an unstable minutia in the red circle. Notice how the different levels of enhancement affect the extraction results. If the strength of the ridge enhancement filter is relatively mild, the crossover ridge structure and the corresponding minutiae are maintained (Fig. 6(c)). On the other hand, the crossover minutiae completely disappear when a strong filter is applied (Fig. 6(e)).

Despite this drawback, however, it is still beneficial to adopt strong filtering since it is capable of consistently extracting stable minutiae even from poor quality images as illustrated by the following example. The image in Fig. 7(b) represents an ideal, manually edited minutiae of the image Fig. 7(a), containing a total of 76 minutiae, 55 of which are stable and 21 are unstable. The image in Fig. 7(d) is an artificially produced poor quality image by covering it with several circular "patches." When a strong filter is applied to the images in Figs. 7(a) and (d), most of the 55 stable minutiae are correctly extracted as shown in Figs. 7(c) and (e), respectively. It should also be noted that this method is especially effective when the area of the overlapping region between the two images is large enough in which a sufficient number of stable minutiae exist.

Thus, filtering strength depends on the operational strategy, requirements, and target image characteristics.

Once fingerprint ridges are suitably enhanced, ▶ fingerprint binarization is then conducted to produce a black and white image and, finally, ▶ fingerprint skeletonization to generate a skeleton image.



Fingerprint Image Enhancement. **Figure 7** Effects of Strong Ridge Enhancement Filter (NIST DB#27 076T). Notes: (**b**) The ideal skeleton image was manually generated so that skeleton curves correctly coincided with the original ridges. Then, minutiae were extracted from the ideal skeleton image. (**c**) The ridge image in (**a**) was automatically enhanced by one of the popular algorithms [9] with a relatively strong enhancement parameter. Then, a skeleton image and minutiae were automatically extracted from that enhanced ridge image. (**e**) The ridge image in (**d**) was automatically enhanced by one of the popular algorithms [9] with a relatively strong enhancement parameter. Then, a skeleton image and minutiae were automatically extracted from that enhanced ridge image.

Summary and Future Improvement

Fingerprint image enhancement is a very effective tool for improving ridge clarity. Undoubtedly, improvement in matching accuracy reported in the past two to three decades can be attributed to innovation in image enhancement techniques. Unfortunately, it is far from true if considered in terms of how close the automated fingerprint recognition got to the ability of human perception. This is because the current techniques that heavily rely on ridge orientation and frequency (and whatever information one can think of) are not capable of perceiving a fingerprint image as a fingerprint but just a collection of gray-scale pixels, and the circumstance has not changed in the course of over 40 years of research. This may change in the future if a leap forward in the computational neuroscience reveals the mechanism of human pattern recognition, but for the time being, a goal pro tempore is probably to find a way to extract information from unmodified gray images to avoid side effects of the image enhancement as far as possible.

Acknowledgments

The sample images are courtesy of the NIST and the University of Bologna [14, 15]. The author would like to thank Amane Yoshida for proofreading and rewriting in English. For more information on the technical details, refer to 'Fingerprint Analysis and Representation' in *Handbook of Fingerprint Recognition* by Maltoni Et al. [2].)

Related Entries

► Fingerprint Classification
► Fingerprint Features
► Fingerprint Image Quality

References

1. ANSI/NIST-ITL 1-2007 Fingerprint Standard – http://fingerprint.nist.gov/standard/index.html
2. Maltoni, D. et al.: Handbook of fingerprint recognition. Springer (2003)
3. Xia, X., O'Gorman, L.: Innovation in fingerprint capture devices. Pattern Recognit. **36**, 361–369 (2003)
4. Hara, M.: Image Density Conversion Method, Image Enhancement Processor, and Program Thereof (USP 20080050030A1 – Pending)
5. Hong, L et al.: Fingerprint image enhancement: Algorithm and performance evaluation. IEEE Trans. Pattern Anal. Mach. Intell. **20**, 777–789(1998)
6. Stock, R.: Automatic fingerprint reading. In: The 1972 Carnahan Conference, on Electronic Crime Countermeasures, April 19–21, 1972
7. Asai, K. et al.: Automatic fingerprint identification. SPIE vol. 182, Imaging Application for Automated Industrial Inspection & Assembly (1979).
8. Capello, R. et al.: Method and apparatus for contextual data enhancement (USP 4,876,726)
9. Kamei, T. et al.: Image filter design for fingerprint enhancement. In:Proceedings International Symposium on Computer Vision, 109–114 (1995).
10. Chikkerur, S. et al.: Fingerprint enhancement using STFT analysis. Pattern Recognit. **40**, 198–211 (2007), 109–114 (1995)
11. Funada, et al.: System and method for processing fingerprint/palmprint image (USP 7,027,626)
12. Hara, M.: System for recognizing fingerprint image, method and program for the same (USP 20070036401A1 – Pending)
13. Paul, A. et al.: A study of image enhancement techniques for fingerprint identification. Proceedings of the IEEE International Conference on Video and Signal Based Surveillance (AVSS'06) (2006)
14. NIST (National Institute of Standard and Technologies) Special Database #27 – http://www.nist.gov/srd/nistsd27.htm
15. FVC2002 Second Fingerprint Verification Competition Database – http://bias.csr.unibo.it/fvc2002/

Fingerprint Image Quality

Elham Tabassi, Patrick Grother
National Institute of Standards and Technology, MD, USA

Synonym

Expected performance or utility of fingerprint image in an automated comparison environment

Definition

The intrinsic characteristic of a biometric signal may be used to determine its suitability for further processing by the biometric system or assess its conformance to preestablished standards. The quality of a biometric

Fingerprint Image Quality. Figure 1 Good quality fingerprint images (**a**) have clear pattern of ridge and valleys; however, poor quality fingerprint images (**b**) do not have easily distinguishable patterns. Poor quality images result in spurious and missed features, thus degrading the performance of the overall system. Poor quality samples can be due to distorted source like abraded skin (**b**), distortion in one or more steps of the process, e.g., capture (residual fingerprints on the platen in (**c**)) or compression, or low character source, the sample may subjectively be assessed as ''good'' quality, but a matcher may not be able to match it to its mate (**d**).

signal is a numerical value (or a vector) that measures this intrinsic attribute. Quality score is a quantitative expression of the utility, or predicted performance of a biometric sample in a comparison environment. This means that finger image quality scores should correlate to the observed false match and ► false non-match rates of the samples.

Introduction

With an increase in the need for reliable identity authentication, biometric recognition systems have been increasingly deployed in several different applications: government applications such as national ID card, border control; and commercial applications, such as physical access control, e-commerce, or mobile phone. Among all biometric modalities, fingerprint recognition is the most widespread due to its permanence and uniqueness [1].

A fingerprint is a pattern of friction ridges on the surface of a fingertip. A good quality fingerprint has distinguishable patterns and features that allow the extraction of features, which are useful for subsequent matching of fingerprint pairs. This viewpoint may be distinct from the human conception of quality. If, for example, an observer sees a fingerprint with clear ridges, low noise, and good contrast then he or she might reasonably say it is of good quality. However, if the image contains few minutiae points then a minutiae-based matcher would underperform. Thus, in the context of automated matching, the term quality should not be used to refer to the fidelity of the sample, but instead to the utility of the sample to an automated system. Figure 1 shows examples of good and poor quality fingerprint images.

Automatically and consistently determining the quality of a given biometric sample for identification and/or verification is a problem with far-reaching ramifications. If one can identify low quality biometric samples, this information can be used to improve the acquisition of new data. This same quality measure can be used to selectively improve an archival biometric database by replacing poor quality biometric samples with better quality samples. Weights for multimodal biometric fusion can be selected to allow better quality biometric samples to dominate the fusion. All of these applications require that the quality of the biometric sample be determined prior to identification or verification. Most of these applications also require that quality of the biometric sample be computed in real-time during data acquisition.

Fingerprint Image Quality

Performance of an automated fingerprint recognition system is greatly affected by the degree of imperfection present in the finger image. Accuracy of current fingerprint recognition systems is high when high-quality samples are being compared [2] (Note that according to Minutia Interoperability Exchange Test 2004 (MINEX04) report, best single finger proprietary fingerprint recognition system performed at 0.0047 false non-match rate at 1% **false match rate**.). However, performance degrades substantially as quality drops.

Although only a small fraction of input data are of poor-quality, the bulk of recognition errors can be attributed to poor-quality samples.

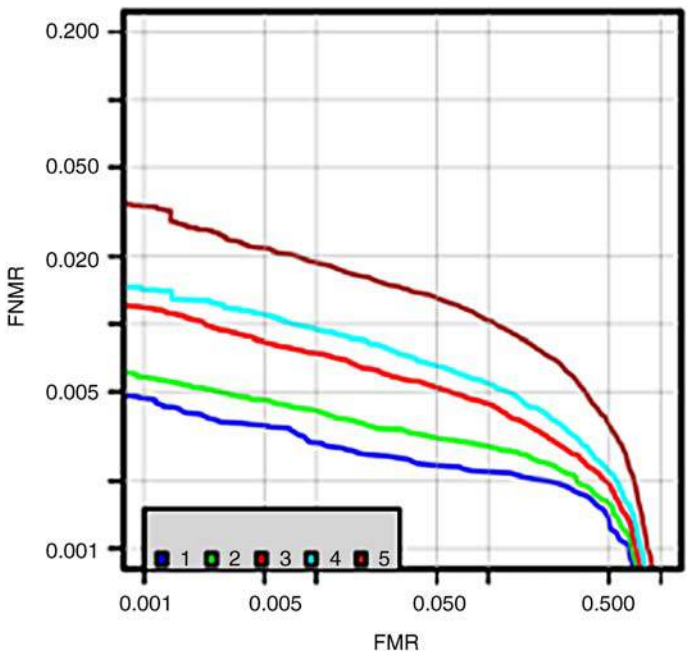
Degradation in fingerprint image quality reduces the amount of identifiable information in a fingerprint. Poor quality images cause spurious and missed features which decrease the likelihood of a correct verification and/or identification, while extremely poor quality samples might be impossible to verify and/or identify. The variation in performance for different quality levels is shown in Fig. 2. The five traces of Detection Error Tradeoff (DET) curves correspond to five different levels of quality as measured by NIST Fingerprint Image Quality (NFIQ) [3, 4]. NFIQ is an integer between 1 and 5 where 1 represents the highest quality and 5 the lowest (unusable) quality.

Several factors affect the quality of fingerprint images: user's skin condition, improper finger placement, scanner limitation or imperfection, impurities on the scanner surface and others. The cause of these imperfections can be classified in four groups: (1) *impairments in the source of* ▶ Biometric characteristics: like scars, blisters, skin conditions such as wet or dry, age, occupation, etc.; (2) *user behavior:* such as improper finger placement, e.g., rotating finger or placing only tip of a finger which cause capturing insufficient area of finger image; (3) *imaging:* e.g., low contrast, distortion, sampling error, insufficient dynamic range, etc.; and (4) *environment:* such as temperature, humidity, or unclean platen.

Fingerprint Image Quality Measures

It is widely accepted that a statement of a biometric sample's quality should be related to its recognition performance. That is, a quality measurement algorithm takes a signal or image, $\mathbf{x}$, and produces a scalar, $q = Q(\mathbf{x})$, which is predictive of error rates associated with the verification or identification of that sample. This predictive value of quality measures may be imperfect but valuable nevertheless. It should be noted that operationally the requirement for a scalar is not necessary: a vector could be stored and could be used. The fact that quality has historically been conceived of as scalar is a widely manifested restriction [5].

International Standards Organization (ISO) has recently established a biometric sample quality draft standard [6], in which quality score of a biometric sample is defined as predicted performance of the



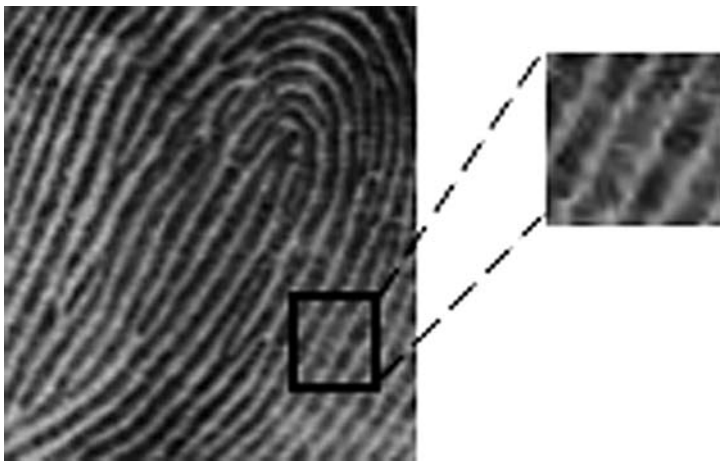
Fingerprint Image Quality. **Figure 2** Quality ranked detection error trade-off characteristics. Five traces correspond to five NFIQ levels. Fingerprint images with NFIQ = 1 (highest quality) cause lower recognition error than images with NFIQ = 5 (lowest quality).

sample in a comparison environment. It considers three components of quality: (1) *character*, which refers to quality of inherent physical features of the source, for example, a fingerprint with a scar has low character; (2) *fidelity*, which is the degree to which a sample is an accurate representation of its source, for example, distortion degrades fidelity; and (3) *utility*, which refers to contribution of a sample to the overall biometric recognition error rates and is related monotonically to the performance of biometric matchers. Character and fidelity of a sample positively or negatively impact the utility of the sample.

There are several fingerprint analysis approaches that gauge character and fidelity of fingerprint images. These measures are then summarized into a scalar (or a vector) quality score that is indicative of utility of the sample. Broadly fingerprint image analysis can be divided into local and global analysis methods [7]. Fingerprint local structure constitutes the main texture-like pattern of ridges and valleys within a local region while valid global structure puts the ridges and valleys into a smooth flow for the entire fingerprint. The quality of a fingerprint image is determined by both its local and global structures. Local feature analysis methods partition an image into nonoverlapping blocks and assign a quality score to each block which indicates the amount of useful information in that block for subsequent matching. Final image quality score can be computed by combining quality scores of the blocks. Global feature analysis examines continuity and uniformity of ridge–valley structure of a fingerprint image in a holistic manner and computes a global measure of fingerprint quality.

Global and local quality measures could be combined to obtain final quality score of a fingerprint image such that the overall quality score is a measure of matchability of the sample in an automated matching process, i.e., the derived quality score should be related to the biometric error rates that is likely to be realized when the sample is matched.

1. *Local Analysis* To locally analyze a fingerprint image, it is divided into grids of blocks (Fig. 3). For each block, local features such as directional flow of ridges are computed which are then summarized into a quality score representing quality of the block. Each block should be large enough to contain sufficient ridge–valley information, at least two ridges per block. For example, for a fingerprint with a resolution of 500 ppi, each block could be 32×32 pixels. An overview of existing local analysis methods follows.
   a. *Orientation certainty field*: A fingerprint image within a small block generally consists of ridges (dark pixels) separated with valley (light pixels) lines along the same orientation. High-quality blocks of a fingerprint image contain consistent ridge (or valley) orientation. Local angle information in each block can be used to compute local features. Lim et al. [8] computed energy concentration along the dominant direction of



Fingerprint Image Quality. **Figure 3** Local analysis consists of partitioning a fingerprint image into small blocks. Local features such as orientation consistency or directional flow are extracted from each block. These features convey information useful for comparison of the image and therefore indicate quality of the block.

ridges by computing the ratio between two eigenvalues of the covariance matrix of a block's gradient vector. It gives an indication of how strong the energy is concentrated along the ridge–valley orientation. Chen et al. [9] measured orientation coherence in each block using gradient of the gray level image.

b. *Ridge–valley structure*: Well-formed and clearly visible ridges are essential to the reliable detection of ridge endings and bifurcations, also known as minutia points. Ridges that are too close or too far apart, or ridges that are unreasonably thick or thin indicate that the finger image may not have captured properly, due to, e.g., pressing too hard or too soft (Fig. 4). Shen et al. [10] applied Gabor filter to image subblocks, to identify blocks with clear repetition of ridge and valley pattern as good quality blocks.
c. *Pixel intensity or Directional contrast*: Region of good quality exhibits high directional contrast, which means that the ridges and the valleys are well separated with regard to gray values. High-quality blocks will exhibit large variance in gray levels while low-quality blocks will show small variance. [11–13] assess quality of each block based on its pixel intensity. Bolle et al. [14] used ratio of directional area to other nondirectional area as a quality measure.
d. *Power Spectrum*: Ridge and valley structure in a high-quality block forms a periodic signal, which can be approximated either by a square wave or a sinusoidal wave with its frequency lie in certain range. In frequency domain, a square wave exhibits a dominant frequency with sideband frequency components (sinc function), and a sinusoidal wave consists of one dominant frequency and minimum components at other nondominant frequencies. Therefore, existence of a dominant frequency component plus its frequency are indicative of high- quality blocks of fingerprint image. Poor quality blocks will not exhibit a dominant frequency or it will be out of the normal range of ridge frequency [12]. Hong et al. [15] modeled the ridge and valley pattern as sine wave, and computed the amplitude, frequency as well as the variance of the sine wave to decide the quality of the fingerprint. Nill and Bouzas [16] propose an objective image quality based on the digital image power of normally acquires scenes. Their system is designed to assess the quality of digital images and can be applied to fingerprint as well.

2. *Global Analysis* A good quality fingerprint exhibit smooth changes in ridge orientation across the entire fingerprint image except when a core or delta point occurs. Ratio of ridge to valley thickness should also be fairly constant throughout the whole image. [8] used local angle information in each block to assess continuity in orientation field between neighboring blocks and uniformity of ridge to valley thickness ratio. Chen et al. [13] computed a block's absolute difference in the orientation angle with its neighboring blocks as a measure of

Fingerprint Image Quality. **Figure 4** Examples of (**a**) good, (**b**) thin, and (**c**) thick ridge structure. (**b**) and (**c**) pose challenge to automated matching system and hence are of lower quality than (**a**).

smoothness of the change in orientation angles among blocks. As mentioned earlier, the ridges of a finger image can be locally approximated by one sine wave with its frequency in a certain range. A region of interest (ROI) of the spectrum is defined as an annular region with radius ranging between the minimum and maximum typical ridge frequency values. For a more robust ridge structure (i.e., the better image quality) the energy will be more concentrated within the ROI. [9] measured the energy concentration in ring-shaped regions of the ROI by employing bandpass filters to extract the energy in each frequency band. Good quality images will have the energy concentrated in few bands while poor quality fingerprints will have a more diffused distribution.

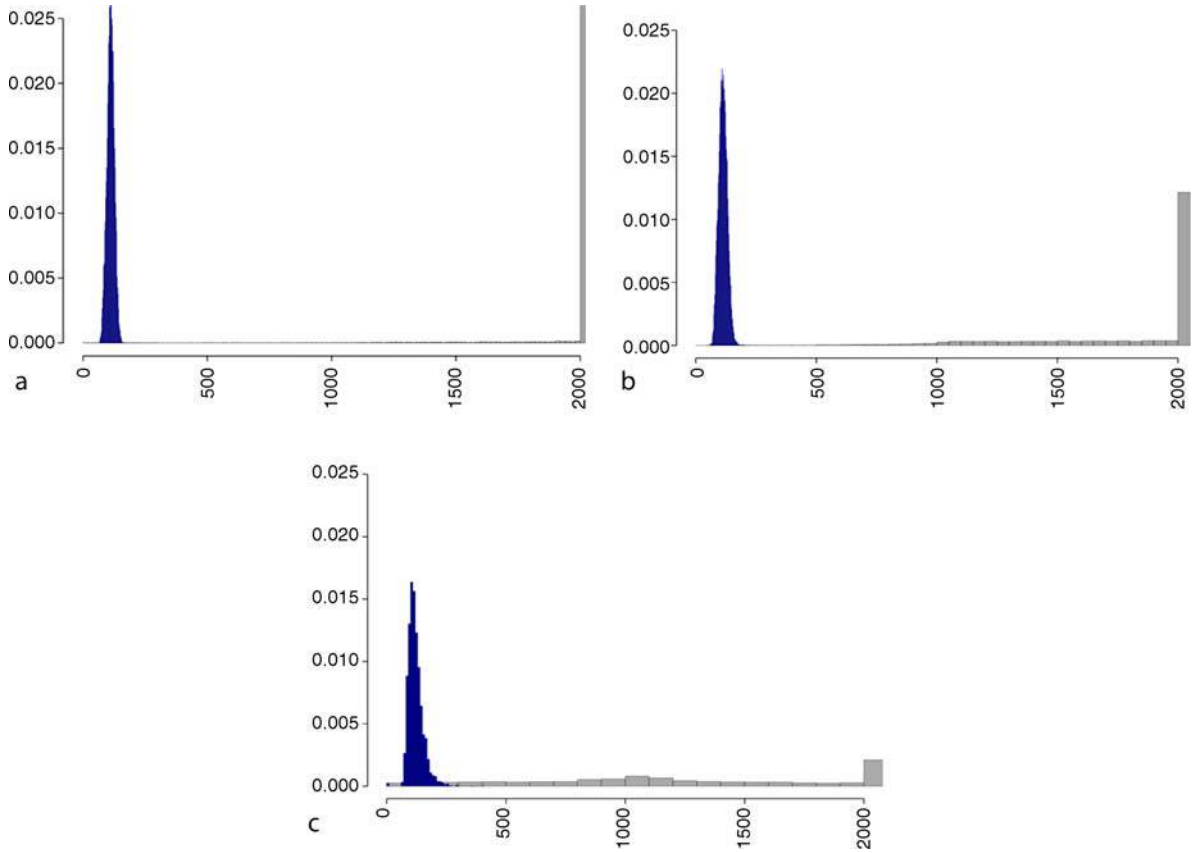
3. *Overall Fingerprint Image Quality: prediction of performance*: It is desirable to combine local and global quality features into one scalar or a vector of quality such that the overall fingerprint image quality is related to the expected false match and false non-match of the image. The summarization can simply be the percentage of blocks classified as "good" or "bad" quality after a local analysis, or more elaborate combination methods such as weighted average of local qualities. For example, higher weights could be assigned to blocks closer to the centroid of a fingerprint since features extracted from blocks near the centroid have more useful and reliable information [9, 11]. Use of a classifier to nonlinearly combine local and global features was first proposed by Tabassi et al. [3, 4]. The method called NIST Fingerprint Image Quality NFIQ [3, 4] was developed to predict how far a genuine score would lie from its impostor distribution and is thus effective at improving false rejections while suppressing false acceptance errors. NFIQ extracts minutia, assigns a quality value to each minutia point, and measures orientation field, pixel intensity, and directional map to compute the following local and global features: number of foreground blocks, number of minutia, number of minutia that have quality value better than certain thresholds, percentage of foreground blocks of excellent, good, fair, and poor quality. A neural network was trained to classify the computed feature vectors into five levels 1–5 where NFIQ = 1 is the best quality and NFIQ = 5 is the lowest quality. Figure 2 shows that the highest recognition performance is achieved for the best quality samples (NFIQ = 1), and samples with lowest quality (NFIQ = 5) have the lowest performance. The plots of Fig. 5 show, respectively, the genuine and impostor distributions for NFIQ values 1 (excellent quality), 3 (average quality), and 5 (poor quality). The overlapping of genuine and impostor for the poorest NFIQ (i.e., NFIQ = 5) means higher recognition errors for that NFIQ level while the almost complete separation of the two distributions for the best quality samples (i.e., NFIQ = 1) indicates lower recognition error. Source code for NFIQ algorithm can be found in [17].

Applications of Biometric Quality Values

This section describes the roles of a sample quality measure in the various contexts of biometric operations. The quality value here is simply a scalar summary of a sample that is taken to be some indicator of matchability. These uses of biometric sample quality are not fingerprint specific and can be generalized to other modalities like face or iris.

1. *Enrollment Phase Quality Assessment* Enrollment is usually a supervised process, and it is common to improve the quality of the final stored sample by acquiring as many samples as are needed to satisfy either an automatic quality measurement algorithm, a human inspector (a kind of quality algorithm), or a matching criterion (by comparison with a second sample acquired during the same session). Our focus on automated systems' needs is warranted regardless of analyses of these other methods, but the authors do contend that naive human judgment will only be as predictive of a matcher's performance as the human visual system is similar to the matching system's internals, and it is not evident that human and computer matching are functionally comparable.

   Specifically, human inspectors may underestimate performance on overtly marginal samples. Certainly human inspectors' judgment may be improved if adequate training on the failure modes and sensitivities of the matcher is given to the

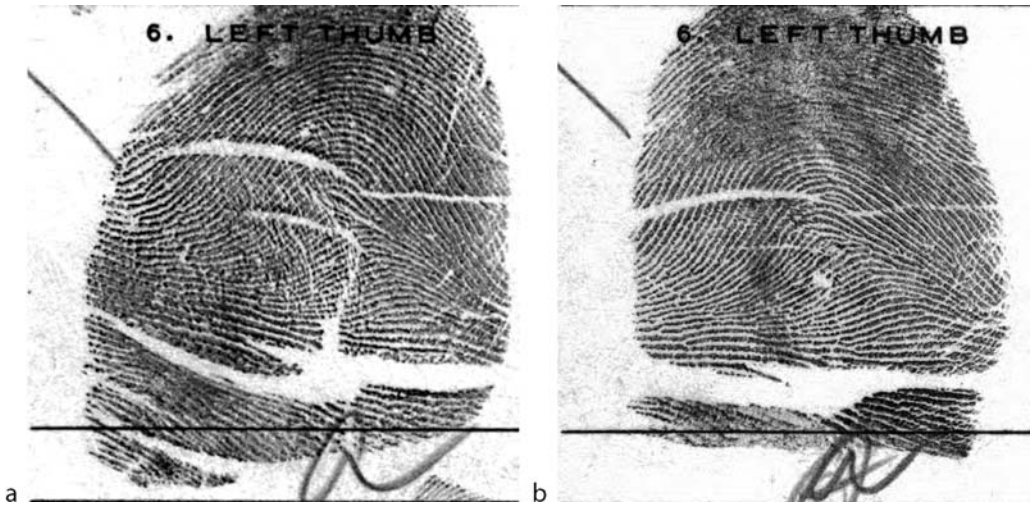


Fingerprint Image Quality. **Figure 5** Probability density of impostor scores is shown in blue and probability density of genuine scores is shown in gray. There is a higher degree of separation between the genuine and impostor distribution for better quality samples as measured by NFIQ. (**a**) Best. (**b**) Middle. (**c**) Worst.

inspector, but this is often prohibitively expensive or time consuming and not scalable. Immediate matching also might not be predictive of performance over time because same-session samples usually produce unrealistically high match scores. For instance, Fig. 6 shows an example of two same-session fingerprint images that were matched successfully by three commercial vendors despite their obvious poor quality.

In any case, by viewing sample acquisition as a measurement and control problem in which the control loop is closed on the quality measure, a system gains a powerful means of improving overall sample quality.

2. *Quality Assurance* Finger image quality assessment algorithms may be used to monitor quality across multiple sites or over time. This is useful to signal possible performance problems ahead of some subsequent matching operation. Quality values may be aggregated and compared with some historical or geographic baselines. Use of quality values in this role has been documented in [18]. The National Institute of Standards and Technology (NIST) has published a technical guidance toward quality summarization [19]. Quality summarization addresses the important issue of enterprise quality-assurance surveying by providing tools on how to combine quality scores of individual samples into one scalar representing quality of the whole database. Such a function would support identification of, e.g., defective sensors, underperforming sites, and seasonal or secular trends.
3. *Verification Quality Assessment* During a verification transaction, quality can be improved by closing an acquire–reacquire loop on either a match-score from comparison of new and enrollment samples or on a quality value generated without matching. Indeed it is common to implement an



Fingerprint Image Quality. **Figure 6** Example of same session captures of single finger that despite their poor quality (NFIQ = 5) were matched correctly by three leading commercial matchers.

"up to three attempts" policy in which a positive match is a de facto statement that the sample was of good quality – even if the individual happens to be an impostor. Depending on the relative computational expenses of sample matching, reacquisition, and quality measurement, the immediate use of a matcher may not be the best solution. The key difference here (as compared with the enrollment-phase) is that quality values of bot h the enrollment and verification samples can be used to predict performance. This two-dimensional problem is distinct from the enrollment case where only one quality value is used.

4. *Identification Quality Assessment* Quality measurement in identification systems is important for at least three reasons. First, many users often do not have an associated enrollment sample. So a one-to-many match will be an inefficient and inconclusive method of stating whether the authentication sample had high quality. Second, in negative identification systems where users with an enrolled sample are motivated to evade detection, quality measurement can be used to detect and prevent submission of samples likely to perform poorly [20], which may help prevent attempts at spoofing or defeating detection. Third, identification is a difficult task: it is imperative to minimize both the false non-match rate (FNMR) *and* the false match rate (FMR). To the extent that consistently high-quality samples will produce high genuine scores, a high matching threshold can be used and this will collaterally reduce FMR. But in large populations FMR becomes dominant, and this raises the question: can a quality apparatus be trained to be directly predictive of false match likelihood?
5. *Differential Processing* Quality measurement algorithms can be used toalter the subsequent processing of a sample. Such conditional activity arecategorized as follows.
   a. *Pre-processing Phase*
      An identification system might apply image restoration algorithms or invoke different ▶ feature extraction algorithms for samples with some discernible quality problem.
   b. *Matching Phase*
      Certain systems may invoke a slower but more powerful matching algorithm when low-quality samples are compared.
   c. *Decision Phase*
      The logic that renders acceptance or rejection decisions may depend on the measured quality of the original samples. This might involve changing a verification system's operating threshold for poor quality samples. For example, in multi-modal biometrics, the relative qualities of samples of the separate modes may be used to augment a fusion process [21, 22].
   d. *Sample Replacement*
      To negate the effects of template aging, a quality measurement may be used to determine whether

a newly acquired sample should replace the enrolled one. An alternative would be to retain both the old and new samples for use in a multi-instance fusion scheme.

e. *Template Update*
Again to address template aging, some systems instead combine old and new sample features. Quality could be used in this process.

Summary

Fingerprint quality measurement is an operationally important task. This paper enumerated ways in which it is useful to compute a quality value from a sample. In all cases the ultimate intention is to improve matching performance. The authors asserted therefore that quality algorithms should be developed to explicitly target matching error rates, and not human perceptions of sample quality. The term quality should not be equated to the acquisition settings of the sample, such as image resolution, dimensions in pixels, grayscale/color bit depth, or number of features. Though such factors may affect sample utility and could contribute to the overall quality score. We reviewed the existing practice of fingerprint local and global analysis. Local and global quality scores could be combined to form a vector of overall finger image quality. However, it is useful, even necessary for some applications, if local and global quality measures are summarized into a scalar which is predictive of error rates associated with the verification or identification of that sample.

Related Entries

► Biometric Sample Quality Standard
► Performance of Quality Measures

References

1. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003)
2. Grother, P., et al.: MINEX: Performance and Interoperability of the INCITS 378 Fingerprint Template. National Institute of Standards and Technology, NISTIR 7296 edn. (2005). http://fingerprint.nist.gov/minex04
3. Tabassi, E., Wilson, C., Watson, C.: Fingerprint Image Quality, NFIQ. National Institute of Standards and Technology, NISTIR 7151 edn. (2004)
4. Tabassi, E., Wilson, C.L.: A novel approach to fingerprint image quality. In: ICIP (2), pp. 37–40 (2005)
5. Tilton, C., et al.: The BioAPI Specification. American National Standards Institute, Inc. (2002)
6. Benini, D., et al.: ISO/IEC 29794-1 Biometric Sample Quality Standard: Framework. JTC1 / SC37 / Working Group 3 (2008). http://isotc.iso.org/isotcportal
7. Alonso-Fernandez, F., Fierrez, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., Fronthaler, H., Kollreider, K., Bigun, J.: A Comparative study of fingerprint image-quality estimation methods. IEEE Trans. Inform. Forens. Secur. **2**, 734–743 (2007)
8. Lim, E., Jiang, X., Yau, W.: Fingerprint image quality and validity analysis. In: IEEE proceedings of International Conference on Image Processing (ICIP), pp. 469–472. New York, USA (2002)
9. Chen, Y., Dass, S.C., Jain, A.K.: Fingerprint quality indices for predicting authentication performance. In: AVBPA, pp. 160–170 (2005)
10. Shen, L., Kot, A.C., Koo, W.M.: Quality measures of fingerprint images. In: AVBPA, pp. 266–271 (2001)
11. Ratha, N., Bolle, R.: Automatic Fingerprint Recognition Systems. Springer, New York (2004)
12. Lim, E., Toh, K.A., Saganthan, P.N., Jiang, X., Yau, W.Y.: Fingerprint image quality analysis. In: ICIP, pp. 1241–1244 (2004)
13. Chen, T.P., Jiang, X., Yau, W.Y.: Fingerprint image quality analysis. In: ICIP, pp. 1253–1256 (2004)
14. Bolle, R., et al.: System and methods for determining the quality of fingerprint images. US Patent 596356 (1999)
15. Hong, L., Wan, Y., Jain, A.K.: Fingerprint image enhancement: algorithm and performance evaluation. IEEE Trans. Pattern Analy. Mach. Intell. **20**(8), 777–789 (1998)
16. Nill, N., Bouzas, B.H.: Objective image quality measure derived from digital image power spectra. Opt. Eng. **31**(4), 813–825 (1992)
17. National Institute of Standards and Technology: NIST Biometric Image Software (NBIS) (2008). http://www.itl.nist.gov/iad/894.03/nigos/nbis.html
18. Ko, T., Krishnan, R.: Monitoring and reporting of fingerprint image quality and match accuracy for a large user application. In: Proceedings of the 33rd Applied Image Pattern Recognition Workshop, pp. 159–164. IEEE Computer Society (2004)
19. Tabassi, E., Grother, P.: Quality Summarization: Recommendations on Enterprise-wide Biometric Quality Summarization. National Institute of Standards and Technology, NISTIR 7244 edn. (2007)
20. Wein, L.M., Baveja, M.: Using fingerprint image quality to improve the identification performance of the u.s. visit program. In: Proceedings of the National Academy of sciences (2005). www.pnas.org/cgi/doi/10.1073/pnas.0407496102
21. Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., Bigun, J.: Discriminative multimodal biometric authentication based on quality measures. Pattern Recogn. **38**(5), 777–779 (2005)
22. Tabassi, E., Quinn, G.W., Grother, P.: When to fuse two biometrics. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR-06. New York (2006). Biometric Workshop

Fingerprint Indexing

George Bebis
University of Nevada, Reno, NV, USA

Synonyms

Continuous classification; Fingerprint retrieval; fingerprint authentication; fingerprint identification

Definition

When matching a query fingerprint to a large fingerprint database for identification purposes, a critical issue is how to narrow down the search space. Indexing provides a mechanism to quickly determine if a query fingerprint is in the database and to retrieve those fingerprints that are most similar with the query, without searching the whole database.

Introduction

Fingerprint matching is one of the most popular and reliable biometric techniques used in automatic personal identification. Typically, fingerprint matching is based on low-level features determined by singularities in the finger ridge pattern known as *minutiae*. To be practical, matching should be robust to translation, rotation, scale, shear, occlusion, and clutter. In this context, matching two fingerprints implies finding a subset of minutiae in the first fingerprint that best match to a subset of minutiae in the second fingerprint through a geometric transformation in an optimal sense.

There are two main applications involving fingerprint matching: *fingerprint authentication* and *fingerprint identification*. While the goal of fingerprint authentication is to verify the identity of a person, the goal of fingerprint identification is to establish the identity of a person. In this case, matching involves comparing a query fingerprint against a database of reference fingerprints to establish the identity of the query. An important issue in fingerprint identification is how to select the most similar fingerprint(s) to the query fingerprint from the fingerprint database. The easiest but least effective way to search a large database is to compare the query fingerprint with each fingerprint in the database. Since usually there is no *a-priori* knowledge of possible correspondences between the query and the reference fingerprints, however, matching can be computationally too expensive, even for a moderate number of reference fingerprints.

A common approach to narrow down the search is by dividing the fingerprint database into smaller sets using *fingerprint classification*. The idea is to match the query fingerprint against fingerprints of the same type only. Although this approach can reduce the number of matches, it is not very effective since fingerprints are unevenly distributed (i.e., more than 90% of the fingerprints belong to only three classes [1]). Several *subclassification* systems have been proposed to address this issue by further dividing some of the classes into more specific categories, however, these systems are much more complex and difficult to implement [1].

A more effective approach to narrow down the search space is to use *indexing*. In principle, indexing can quickly determine if a query fingerprint is in the database and to retrieve those reference fingerprints which are most similar to the query fingerprint, without searching the whole database. Therefore, methods based on indexing are less dependent on the size of the database. The main idea is to assign an index value to each fingerprint and match the query against those reference fingerprints having comparable indices only. Indexing methods have been very popular in computer vision for searching large databases of models in object recognition [2–4]. Therefore, many indexing schemes for finger identification have their roots in object recognition.

How Indexing Works

Indexing is a mechanism which, when provided with a key value, can rapidly access some associated data. Thus, instead of searching the space of all possible matches and explicitly rejecting invalid ones, indexing inverts the process so that only the most feasible matches are considered for matching. In essence, indexing serves as a "filtering" step which allows to verify a query fingerprint against the most similar fingerprints in the database only. To implement indexing, certain information about the reference fingerprints is prestored in an index structure. During identification, the index structure is accessed efficiently to narrow down the search.

Typically, a single index can be computed from the whole fingerprint or multiple indices can be computed from groups of local features. Using a single index, fingerprints are mapped to numerical vectors in a high-dimensional space through a similarity-preserving transformation. During identification, the query fingerprint is compared against those reference fingerprints which are close to the query in the multidimensional space. This approach, also known as *continuous classification* [5], is in essence a classification approach, however, the classes are not disjoint. Commonly, the orientation image is used in the mapping transformation, however, different transformations and distance measures have been proposed [5–8].

Using groups of local features, reference fingerprints are represented redundantly in the database by computing a separate index for each group of features and making an entry for each index [9–11]. This kind of redundancy provides robustness during identification by allowing the retrieval of reference fingerprints that match the query fingerprint only partially. Specifically, for each reference fingerprint, groups of features are extracted and an index is constructed from each group. The indexed locations are filled with entries containing information about the reference fingerprints. At a minimum, each entry contains information about the identity of the reference fingerprint and the group of features that generated the index.

During identification, the information stored in the index structure is used to quickly eliminate noncompatible matches between the query and the reference fingerprints. To reduce the number of false matches, geometric constraints can be used [11]. The reference fingerprints listed in the indexed locations are collected into a list of candidate fingerprints and the most often indexed fingerprints are selected for further verification. Verification works by computing the transformation between the candidate fingerprints and the query. Then, the candidate fingerprints are aligned with the query and their similarity to the query is estimated by finding the percentage of candidate features that have been aligned with query features.

An Example

Here is an example, based on [9, 10], to illustrate the use of indexing for fingerprint identification. In this example, matching a pair of minutiae sets is performed by comparing minutiae triangles, formed by minutiae triplets, using geometric invariant features. In general, a pair of corresponding minutiae triangles provides enough information to compute a geometric transformation (e.g., similarity or affine) that potentially aligns the minutiae sets. To compute good alignments, voting can be applied in the transformation space to find transformations that are supported by many minutiae triangles [9]. A number of hypothetical transformations is obtained by considering transformations that have received a high number of votes. Each hypothetical transformation is then explicitly verified by counting the number of aligned minutiae.

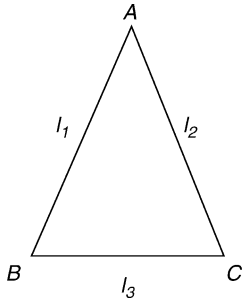
The indexing mechanism used in this example is based on geometric hashing [2]. Specifically, given a triplet of minutiae, three geometric invariants can be computed by considering the triangle formed by the minutiae triplet. The geometric invariants are based on the sides and angles of the minutiae triangle, as shown in Fig. 1, and remain unchanged under similarity transformations (i.e., translation, rotation, and scale). First, the sides of the triangle are sorted to avoid considering all possible orderings:

$$l_1 \leq l_2 \leq l_3$$

Then, we compute the following geometric invariants:

$$0 \leq \frac{l_1}{l_3} \leq 1$$
$$0 \leq \frac{l_2}{l_3} \leq 1$$
$$-1 \leq cos(A) \leq 1$$

where A is the angle between the smallest two sides. To compute an integer index, a simple hash function is applied on the geometric invariants which involves linear scaling followed by quantization. For each index,



Fingerprint Indexing. Figure 1 A minutiae triangle defined by a minutiae triplet (A,B,C).

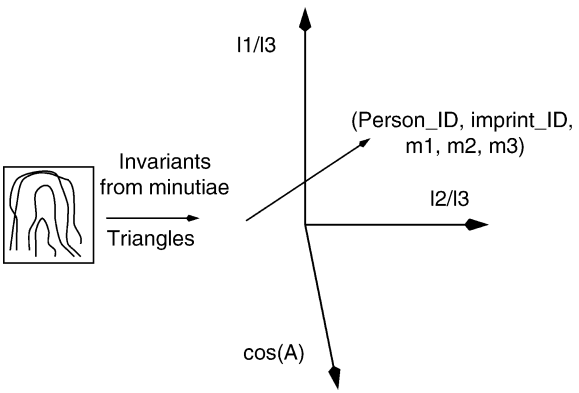
information is stored about the fingerprint and the minutiae triangle in a hash table. Each entry stored in the hash table has the following format:

$$(person_ID, print_ID, m_1, m_2, m_3)$$

where *person_ID* corresponds to the identity of the person whose fingerprint is considered, *print_ID* is an identification code for the particular fingerprint of that person, and m_i are the (x,y) coordinates of the m_i minutia in the triangle. Figure 2 illustrates the indexing step.

During identification, each index generated by the query fingerprint is used to retrieve all reference fingerprints stored in the hash table under the same index. For each minutiae triangle, the lengths of the sides are computed, sorted in ascending order, and the geometric invariants are computed as before. Then, the invariants are scaled and quantized in the same manner. The resulting index is used to extract all entries from the database stored at the same index table location. To account for noise, entries stored in a small neighborhood around the indexed location could be also retrieved.

Several indexing-based approaches accumulate evidence about reference fingerprints by casting a vote for every entry stored in the indexed locations and by "histograming" the entries to pick the ones which have received a high number of votes. However, this approach takes into consideration only the number of votes received by a particular entry and not whether these votes are consistent among themselves. To introduce a measure of coherence, voting in the transformation space has been proposed [9]. The idea is simply to consider transformations which form large clusters in the transformation space.



Fingerprint Indexing. Figure 2 Pre-storing information about the reference fingerprints using indexing.

Each of the entries retrieved from the index table represents a hypothesized correspondence between minutiae triplets in the query and a reference fingerprint. Given this information, the transformation that best maps the query triplet to the reference triplet is computed. The computed transformation parameters are binned and, along with the *person_ID* and *imprint_ID*, form a key that indexes another data structure used for evidence accumulation. An eight-dimensional integer array is used to store the number of votes in the transformation space (i.e., six dimensions for the parameters of the transformation, one for the *person_ID* and one for the *imprint_ID*).

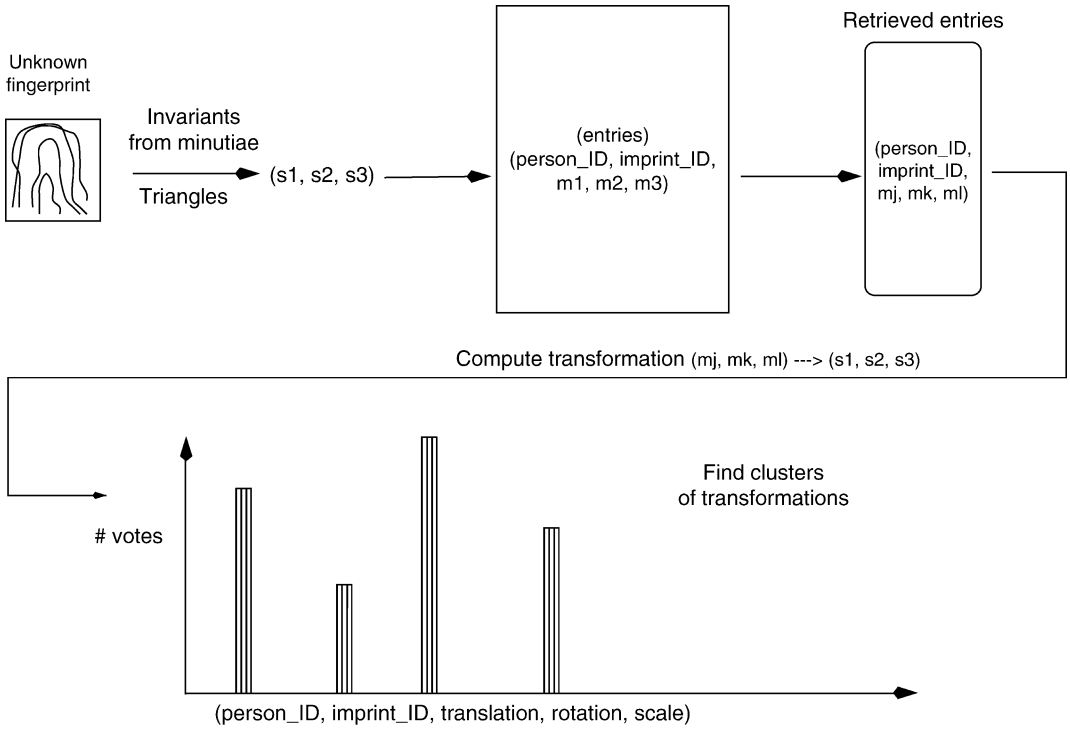
If a large number of minutiae points can be brought into correspondence by a transformation, then all the indices generated by the triangles formed by those minutiae points will yield close transformation parameters. Hence, a larger number of votes for a correct match will be accumulated. Although there might be a number of random correspondences between minutiae triplets in the query fingerprint and some arbitrary reference fingerprints, the likelihood of a number of consistent transformation parameters being generated by random correspondences is small, and the verification step will eliminate most of them. Figure 3 illustrates these identification procedure based on indexing.

Practical Issues

Several important issues must be considered while employing indexing for fingerprint identification including: index construction, index selectivity, storage requirements, indexing mechanism, performance analysis, and error analysis. Each of them briefly discussed in the following section.

Index Construction

As illustrated in the earlier example, each index is typically constructed from groups of local features, such as minutiae triplets. In general, index construction should be based on features that are robust to fingerprint distortions, occlusion, and noise [11]. To reduce storage requirements, the computation of the index is based on geometric invariant features,



Fingerprint Indexing. **Figure 3** Illustration of the identification step using indexing.

that is, features that remain unchanged under certain geometric transformations. In the earlier example, we used length ratios and angles which are invariant to similarity transformations (i.e., translation, rotation, and scale). Other geometric invariant features include ridge count, triangle handedness, triangle type, triangle direction, and maximum side, minutiae density, and various ridge invariants [9, 11–13].

Index Selectivity

Although indexing is an attractive approach, very often it becomes less effective because of limited index selectivity. The issue of index selectivity relates to the discrimination power of the features considered for indexing. Features with low discrimination power give rise to very similar indices (i.e., low index selectivity). As a result, a large number of hypothetical matches can be generated during identification, making indexing ineffective. One way to deal with this problem is to increase the index dimensionality using larger groups of features, however, this would also increase memory requirements since the number of groups increases exponentially with group size. Alternatively, additional information can be computed from each group and added to the index to increase its dimensionality. For example, the FLASH algorithm, introduced in [3] for object recognition and adopted in [9] for fingerprint identification, computes a nine-dimensional index from minutiae triangles. It should be mentioned that although this is an effective approach, it increases time requirements and raises the issue of computing the additional features fast and reliably. Recent studies using high-dimensional indices include [11] and [12].

Storage Requirements

Indexing methods have high storage requirements as they trade space for speed. For example, the number of entries to be indexed using minutiae triplets is of the order of $O(N^3)$ where N is the average number of fingerprint minutiae. If M is the number of fingerprints to be indexed, the total space requirements is of the order of $O(MN^3)$. To reduce storage requirements, geometric constraints can be used to limit the

number of minutiae triangles considered for indexing [9]. Alternatively, a unique topological structure can be associated with the fingerprint minutiae using the Delaunay triangulation [10, 13]. This approach considers only $O(N)$ minutiae triangles for indexing leading to significant memory savings and faster identification. A problem with this approach is that it is sensitive to noise and distortions (e.g, introduced by missing or spurious minutiae), however, both noise and distortion have only a local effect on the triangulation. Nevertheless, hierarchical matching schemes have been proposed to deal with these issues [14].

Indexing Mechanism

Hashing has been the most common indexing mechanism used both in fingerprint identification and object recognition. Hashing performs a range search, retrieving all points within a certain distance from the query point. However, the highest probability hypotheses can be discovered by observing just a few of the closest neighbors. Hashing is not efficient for nearest-neighbor search in high dimensions since it requires time exponential in the dimension of the space (i.e., the nearest neighbors might not lie in the same hash bin as the query point, but in one of the many adjacent bins). Moreover, "good" hash functions are required for distributing the data uniformly [15, 16]. In general, more effective indexing mechanisms can be employed, such as kd-trees [17], to retrieve only the k nearest points.

Kd-trees are data structures used to divide the data into hypercubes containing equal numbers of data. When a query point is presented, the boundaries between the hypercubes are used as decisions to discover the hypercube that contains the query point, and the data in this hypercube will be close matches. To guarantee that the matches in the hypercube containing the query point are in fact closer to the query point than data lying just over the boundary of the hypercube, it is necessary to examine neighboring hypercubes. This can make search quite slow. To deal with this issue, approximate nearest-neighbor schemes can be used which maintain good performance even in quite high dimensions (i.e., 10–20) and large number of data [18, 19]. These algorithms have been demonstrated to uncover the exact nearest neighbor a high percentage of the time and a very close neighbor in the remaining cases.

Performance Analysis

To analyze the performance of indexing schemes, it is typical to use identification rate versus ▶ penetration rate graphs. The ratio of fingerprints retrieved over the size of the database. These graphs show the identification rate achieved by varying the penetration rate. Typically, a low penetration rate with a high identification rate is desirable. Close to 99% identification accuracy with only 5% penetration rate is reported in [12] on DB1 from FVC2002. Alternative measures include the ▶ Correct Index Power (CIP) and the ▶ Correct Reject Power (CRP) [11]. CIP is defined as the number of correctly retrieved fingerprints over the size of the database while CRP is defined as the ratio of correctly rejected reference fingerprints over the number of query images not having a corresponding fingerprint in the database. Using the NIST-4 special database and extrapolating the results from 2,000 images to 30,000 images, Bhanu et al. [11] report a CIP rate of 50% using the top 100 candidate matches (i.e., 0.33% penetration rate). Using a smaller database (i.e., 400 image pairs) and assuming the top candidate match, they report a CIP rate of 96.2% for good quality images, 85.5% for fair quality images, and 83.3% for low quality images. Using the top five candidate matches, the CIP rate increases to 100, 99.2 and 98% correspondingly. The CRP rate reported using 200 query fingerprints not in the database was 100%.

Error Analysis

In the noiseless case, each indexed location will contain exactly the set of reference groups compatible with the query group used to access the index structure. In practice, however, several different sources of error must be taken into consideration to improve robustness. The most common source of errors is from the feature extraction step. Using minutiae triplets, for example, errors in the localization of the minutiae can lead to errors in the computation of the geometric invariants and, as a result, to errors in the computation of the indices. In this case, the correct entries will not be found in the indexed location but in a neighborhood around it. Several studies have considered the effect of localization errors on indexing performance

for object recognition [20]. Other studies model localization errors probabilistically in order to estimate the appropriate neighborhood size to retrieve the correct entries [15].

Summary

Indexing is an attractive method for reducing the number of matches when comparing a query fingerprint with a fingerprint database for identification purposes. This chapter reviewed the main concepts behind fingerprint indexing and discussed several critical issues to be addressed in practice.

Related Entries

- ▶ Fingerprint Authentication
- ▶ Fingerprint Classification
- ▶ Fingerprint Identification
- ▶ Fingerprint Matching

References

1. Maltoni, D., Maio, D., Jain, A., Prabhakar, S.: Handbook on fingerprint recognition. Springer, Berlin (2003)
2. Lamdan, Y., Schwartz, J., Wolfson, H.: Affine invariant model-based object recognition. IEEE Trans. Robot. Automat. **6**(5), 578–589 (1990)
3. Califano, A., Mohan, R.: Multidimensional indexing for recognizing visual shapes. IEEE Trans. Pattern Analy. Mach. Intell. **16**(4), 373–392 (1994)
4. Bebis, G., Georgiopoulos, M., Shah, M., da Vitoria Lobo, N.: Indexing based on algebraic functions of views. Comput. Vision Image Understand. **72**, 360–378 (1998)
5. Lumini, A., Maio, D., Maltoni, D.: Continuous versus exclusive classification for fingerprint retrieval. Pattern Recogn. Lett. **18**(10), 1027–1034 (1997)
6. Cappelli, R., Lumini, A., Maio, D., Maltoni, D.: Fingerprint classification by directional image partitioning. IEEE Trans. Pattern Analy. Mach. Intell. **21**(5), 402–421 (1999)
7. Cappelli, R., Maio, D., Maltoni, D.: Multispace kl for pattern representation and classification. IEEE Trans. Pattern Analy. Mach. Intell. **23**(9), 977–996 (2001)
8. Li, J., Yau, W.Y., Wang, H.: Fingerprint indexing based on symmetrical measurement. Int. Conf. Pattern Recogn. **1**, 1038–1041 (2006)
9. Germain, R., Califano, A., Colville, S.: Fingerprint matching using transformation parameter clustering. IEEE Computational Science and Engineering **4**(4), 42–49 (1997)
10. Bebis, G., Deaconu, T., Georgiopoulos, M.: Fingerprint identification using delaunay triangulation. In: IEEE, International Symposium on Information, Intelligence, and Systems, pp. 452–459 (1999)
11. Bhanu, B., Tan, X.: Fingerprint indexing based on novel features of minutiae triplets. IEEE Trans. Pattern Analy. Mach. Intell. **25**(5), 616–622 (2003)
12. Feng, J., Cai, A.: Fingerprint indexing using ridge invariants. Int. Conf. Pattern Recogn. **4**, 433–436 (2006)
13. Ross, A., Mukherjee, R.: Augmenting ridge curves with minutiae triplets for fingerprint indexing. In: Prabhakar S., Ross A. (eds.) SPIE Defense and Security Symposium (Biometric Technology for Human Identification IV, vol. 6539) (2006)
14. Uz, T., Bebis, G., Erol, A., Prabhakar, S.: Minutiae-based template synthesis and matching using hierarchical delaunay triangulation. In: IEEE International Conference on Biometrics: Theory, Applications and Systems (2007)
15. Wolfson, H., Rigoutsos, I.: Geometric hashing: An overview. IEEE Comput. Sci. Eng. **4**(4), 10–21 (1997)
16. Bebis, G., Georgiopoulos, M., La Vitoria Lobo, N.: Using self-organizing maps to learn geometric hashing functions for model-based object recognition. IEEE Trans. Neural Networks **9**(3), 560–570 (1998)
17. Li, W., Bebis, G., Bourbakis, N.: Integrating algebraic functions of views with indexing and learning for 3D object recognition. IEEE Workshop on Learning in Computer Vision and Pattern Recognition (2004)
18. Nene, S., Nayar, S.: Closest point search in high dimensions. In: Computer Vision and Pattern Recognition Conference, pp. 859–865 (1998)
19. Beis, J., Lowe, D.: Shape indexing using approximate nearest-neighbor search in high-dimensional spaces. In: Computer Vision and Pattern Recognition Conference. pp. 1000–1006 (1997)
20. Grimson, W., Huttenlocher, D., Jacobs, D.: A study of affine matching with bounded sensor error. Int. J. Comput. Vision **13**(1), 7–32 (1994)

Fingerprint Individuality

Fingerprint Individuality is the study of the extent of which different fingerprints tend to match with each other. It is the most important measure to be ascertained when fingerprint evidence is presented in court as it reflects the uncertainty with the decision of the expert.

- ▶ Individuality of Fingerprints

Fingerprint Matching, Automatic

Jie Tian, Yangyang Zhang, Kai Cao
Center for Biometrics and Security Research & The Key Laboratory of Complex System and Intelligence Science Chinese Academy of Sciences, Institute of Automation Zhingguancun Donglu, Beijing, China

Synonyms

Fingerprint comparing; Automatic

Definition

In contrast to manual fingerprint matching, automatic fingerprint matching can be efficiently operated on a computing machine following a series of preset procedures. Automatic matching compares two given fingerprint templates (raw images or extracted features) and returns their similarity score (in a continuous range) or a binary decision (matched/non-matched).

Introduction

With the increasing expansion of large-scale databases, manual fingerprint matching cannot satisfy the demand of efficiency in many applications. Automatic fingerprint matching simulates how human experts compare the fingerprints to measure the similarity between two given fingerprint templates or to determine whether they come from the same finger [1]. For most fingerprint matching procedures, experts calculate the similarity score of two templates and give the final judgment with a preset threshold. If the score exceeds the threshold, the compared templates are considered matched, otherwise they are non-matched. The templates are the representation of fingerprints, comprising extracted features or the raw images in case of no extraction. The features can be categorized into two kinds: local features (minutiae, pores) and global features (compressed raw fingerprint, ridge pattern, orientation and curvature map).

Fingerprint matching is one of the most important stages in ▶ Automatic Fingerprint Identification System (AFIS). It is really difficult to match the different impressions of the same finger and find the corresponding features reliably because of the following interferential factors. First, there are several kinds of transformation between two impressions, including linear transformation (translation, rotation, and scale) and non-linear distortion. The translation and rotation is caused by the differential finger placement with respect to the sensor surface during different acquisitions, which may result in a partially overlapped area. If the impressions are captured by different sensors with different resolutions, there exists scale variation in the transformation space. The non-linear distortion of fingerprints is inevitable because the capture is a process of mapping a three-dimensional finger to a two-dimensional impression. The pattern of distortion is firstly determined by finger pressure, finger condition, and the characteristics of sensors. Secondly, the quality of raw fingerprints are also influenced by the noise (fingerprint residues from the previous capture), skin condition (dryness, grease, skin disease), and the capture environment (humidity, temperature). Figure 1 displays three examples of these interferential factors in fingerprint matching. In addition, the algorithms of fingerprint enhancement and feature extraction are imperfect and often introduce some mistakes into the extracted features. Errors may be made and accumulated during each of the foregoing stages (orientation estimation, singular points detection and minutiae extraction). These objective factors are likely to generate spurious features or miss genuine features. All the above variations may make the templates from the same finger appear quite different, sometimes more severely than the similar templates from different fingers. Many fingerprint matching algorithms have been proposed in the scientific literature. Most of these algorithms are proved successful when dealing with good-quality fingerprints. However, fingerprint matching is still a challenging task due to the difficulty in matching low-quality, partial, or large-distorted fingerprints.

There have been a series of strategies to cope with the transformation between two fingerprints. In most of typical fingerprint matching processes, alignment is utilized to estimate the optimum linear transformation between two fingerprints. It rotates and translates one of the compared templates in order to make its features mostly overlap the corresponding features in another template. To achieve the optimum feature-pairing requires correctly calculating the parameters of translation and rotation. Note that scale has to be taken

Fingerprint Matching, Automatic. Figure 1 Three examples of these interferential factors in fingerprint matching. (**a**) a pair of fingerprints with large translation and rotation; (**b**) a pair of poor-quality and partially overlapped fingerprints; (**c**) a pair of large-distorted fingerprints [22]. While the corresponding minutiae in blue rectangle are overlapped, the maximal distance of corresponding minutiae in red ellipse is above 100 pixels.

into account when the resolutions of fingerprints vary. Previous researches [2] prove that the performance of the matcher drastically decreased when the compared fingerprints originated from sensors with different resolutions. Fingerprint alignment is certainly an important but time-consuming stage. Therefore, some algorithms [3] attempt to avoid this stage in fingerprint matching. For instance, experts construct

local feature structures invariant to the linear transformation for matching without priory global alignment. Such matching algorithms ignore the global relationship among local features and therefore may lose part of the discriminating information. On the other hand, nonlinear distortion is universal during fingerprint acquisition, so it is needed to develop fingerprint matchers tolerant of the distortion. Some methods [4, 5] allow corresponding features to alter in the predetermined range (tolerance box). Others [6, 7] adopt local feature structures for matching because distortion affects to a lesser degree local areas. Few developers [8, 9] introduce an appropriate model to recover the distortion prior to matching. In general, tolerating more transformations may increase the successful percentage of not only ► genuine matching but ► imposter matching. When designing the matching algorithms, the degree of tolerance needs careful evaluation. Based on the calculated transformations, the correspondences between features can be established through the optimization methods.

Classification

Because fingerprint matching algorithms rely heavily on the stored features in the templates, they can be coarsely classified into three categories in terms of the selection of features:

- Local feature-based matching: The most popular local feature is minutia, which was earliest used in fingerprint matching technologies [5]. Minutiae features are extracted and stored in the templates as sets of points in the two-dimensional plane. They are usually described by the location, orientation, type, and other information in the neighborhood region. Most common minutiae matching are addressed as a point pattern matching problems and many approaches can be applied. Furthermore, several adjacent minutiae are constructed as local structures in various forms of minutiae, such as simplex [10] triangle [6] and so on.
- With the advent of high-resolution fingerprint sensors, more precise local features (pores and ridge contours [11]) are employed in fingerprint matching to satisfy the growing demand and requirements for accuracy. These algorithms usually align two different templates to establish the correspondences between two sets of local features and calculate the similarity score combining all the matched features. Compared to other fingerprint features, local features have several advantages in terms of the template size and its discriminability, but they have inevitable drawbacks in practical usage. Sometimes it is difficult to exactly obtain local features due to its sensitivity to the fingerprint quality and capture area, which seriously degrades the performance.
- Global feature-based matching: The global features represent the fingerprint in a global perspective, many of which are more continuous and smooth everywhere except in some special regions. For poor-quality or partial fingerprints, global features can be extracted more reliably. It is too space-and time-consuming to directly store and compare the map/field of features pixel by pixel. To reduce the template size and simplify the matching, features can be approached with appropriate models and stored as a series of parameters. Global feature-based matching [12] overlaps two given templates with different transformation parameters and estimates the similarity score between the corresponding cells. Compared to local features, the global features have less distinctness, so they are often exploited together with other features or in the preprocessing stage of fingerprint matching.
- Combined feature-based matching: Since the local and global features are somewhat independent and capture contemporary information, it is reasonable to improve the discriminating ability of matching by fusing features. The approaches in this category [13, 14] combine the local and global features in the matching stage with available feature-level fusing strategies. The combination can reinforce the individuality of fingerprints and improve the performance for fingerprint systems on large-scale databases.
- How to select features is pivotal for the effect of feature combination. It is proved that combining the irrelative features will bring the most obvious improvement of accuracy or efficiency. On the other hand, fusing local and global features may result in additional time or memory cost, so the appropriate hierarchical strategy can be utilized to reduce resource consumption. For instance, due to the complexity of alignment, two fingerprints can

be pre-aligned by the modeled orientation field (global feature). Then the similarity score is calculated based on the minutiae (local feature). Pre-alignment is more efficient while matching on the large-scale database.

Performance Evaluation

Performance evaluation is necessary for understanding the limitations and advantages of a fingerprint matching algorithm and addressing its appropriate applications. The performance can be evaluated from different aspects: accuracy, resource consumption, scalability and sensor interoperability.

- The accuracy of matching is evaluated based on the distribution of similarity score in genuine and impostor matching. Genuine matching compares two fingerprint templates from the same finger, whereas impostor matching is for two fingerprint templates from different fingers. The overall accuracy can be illustrated by Receiver Operation Characteristics (ROC) curve, which shows the dependence of False Non-match Rate (FNMR) on False Match Rate (FMR) at all thresholds. A series of indicators are adopted to quantify the accuracy containing Equal Error Rate (EER – the point where FNMR and FMR yield the same value), FMR100 (the lowest FNMR for FMR <= 1%), FMR1000 (the lowest FNMR for FMR <= 0.1%) ZEROFMR (the lowest FNMR for FMR <= 0%), and ZEROFNMR (the lowest FMR for FNMR <= 0%) [15, 16]. Accuracy usually attracts most of the attention in common applications, but the algorithms cannot just be characterized by these indicators.
- Resource consumption can be measured through three aspects: the amount of storage, time, and memory required by the algorithms. The storage cost is measured by the average/maximum size of template for each database. The efficiency is indicated by the average/maximum time in genuine/imposter matching and the memory requirement is measured by the average/maximum size of allocated memory in genuine/imposter matching. The variation of the indicators through the whole database reflects the stability of the tested algorithms.
- The scalability reflects the degradation of the accuracy with the growing scale of database, which is available in one-to-many matching. It should be evaluated on different-scale databases through observing the relationship between the aforesaid indicators and the scale of database. The international competition FpVTE2003 [17] adopted three different-scale databases to evaluate the scalability of the tested algorithm.
- Sensor interoperability denotes the ability to handle the templates obtained from different sensors [18]. The features in templates are sensitive to different characteristics of multiple sensors in a fingerprint system. The comparison of measures (accuracy, efficiency) between *intra-sensor* matching (comparing templates from the same sensor) and *inter-sensor* matching (comparing templates from different sensors) somewhat reflects the interoperability of the matching approach. Research on sensor interoperability is at its fledgling stage and so far there have been no authorized databases or indicators for quantified evaluation.

Performance during evaluation is relative to many objective conditions. Accuracy is influenced by both the characteristics of database (size, average quality, distortion) and the testing ▶ protocol, while resource consumption relies on the hardware capability, so it is meaningless to evaluate the matching approach without considering these conditions. An authentic evaluation should be conducted on the databases that have independent training/testing parts and sufficient fingerprints, and calculate the statistical indicators with a reasonable protocol. Because the above indicators are statistical results, it should be reported how believable the evaluation of these statistics really are. The problem can be addressed by computing the confidence intervals on the distribution of these values [19]. The accuracy of these confidence interval estimates is ascertained by both correct estimation strategies and correct dataset sampling.

The comparison of performance among various matching algorithms is always a controversy. Different algorithms have different advantages and disadvantages; therefore it is unfair to directly conclude one better than the other. Some research displays experiment results conducted on the proprietary databases using different protocols. This makes it difficult to compare the performance fairly. Evaluating and comparing these indicators among different algorithms is required to operate on the

same public databases with the same authoritative protocol and testing environment, such as FVC and FpVTE.

Application

In AFIS, the fingerprint matching can be applied in two distinct models: verification and identification. The verification model is a one-to-one matching (1:1) in which a user states his/her identity by means of an ID and proves it with a fingerprint. A new fingerprint sample taken from the user is compared with the user's previously registered or stored fingerprints. The comparison only occurs once between the input fingerprint image and the selected sample from the database following the claim of the user. If the fingerprints are successfully matched, the user is verified as who he/she is claiming to be, and granted all the privileges and access of the stated user. On the contrary, the identification process is a one-to-many matching (1:N), in which a user need not claim his or her identity. A new impression is taken from the user and compared to the existing fingerprints of registered or stored users in the databases. The identification can be implemented with a sequence of verification between the input template and the query templates in the database. Fingerprint identification requires searching the database for a matched template or several candidates, which is a process more complex than verification. Although satisfactory performances have been reported for fingerprint verification, both the efficiency and accuracy of identification deteriorate seriously by simple extension of a 1:1 verification procedure to a 1:N identification system. It is still necessary to improve the performance of fingerprint matching in the large-scale fingerprint database. Fingerprint classification and indexing techniques are proved effective to narrow down the searching space of verification, which will speed up the identifying process.

Different kinds of applications focus on different requirements. For the same algorithm, the matching threshold can be modulated or other parameters configured to realize trade-off among these performance indicators. There exists a strict relationship between accuracy and resource consumption, FMR and FNMR of each algorithm. For instance, both FMR and FNMR are actually the functions of matching threshold. The decrease in value makes the algorithm more tolerant to the transformations (lower FNMR), but increases the possibility of incorrectly matching two templates from different fingers (higher FMR). Contrarily, if the value increases, the algorithm performs with higher FNMR and lower FMR. According to the given application, the threshold is carefully chosen as suitable for the special requirement. It is difficult to develop a matching approach omnipotent in every scenario, therefore s different applications may at times need different algorithms. The embedded applications (mobile phone, identity card) emphasize limited resources and put significant strain on the recognition reliability, because high performance fingerprint matching approaches tend to be computationally intensive. In this case, we tend to adopt these algorithms with lower resources consumption. In contrast, the resource-unlimited applications equipped with adequate resources attach more importance to accuracy rather than the computation and storage expenditure. For instance, fingerprint matching in network security operates on the distributed computer system with a "Trustworthy Authority + Remote Client" mode, where extreme accuracy is the most crucial target. In these situations, we choose the algorithms that have more accuracy despite of the possible computational complexity.

Summary

Recently, there have been great advances in the research on automatic fingerprint matching. However, the various applications of AFIS in personal identification desire further improvement of the performance of matching algorithms. Recent research demonstrates that fusion in different levels (feature, score, decision) is effective in improving the performance in many aspects, attracting increasing interest. Besides the features, the fusion of multiple independent matchers [20, 21] is likely to ameliorate the accuracy of fingerprint matching.

Related Entries

- ▶ Fingerprint Classification
- ▶ Fingerprint Clustering
- ▶ Fingerprint Matching, Manual
- ▶ Fingerprint Recognition
- ▶ Fingerprint Templates

References

l. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, Berlin (2003)
2. Ross, A., Jain, A.K.: Biometric sensor interoperability: a case study in fingerprints, BioAW 2004. LNCS **3087**, 134–145 (2004)
3. Bazen, A.M., Gerez, S.H.: An intrinsic coordinate system for fingerprint matching. In: Third International Conference on Audio- and Video-Based Biometric Person Authentication, Halmstad, Sweden (2001)
4. Luo, X.P., Tian, J., Wu, Y.: A Minutia matching algorithm in fingerprint verification. Fifteenth ICPR **4**, 833–836 (2000)
5. Jain, A.K., Lin, H., Bolle, R.: On-line fingerprint verification. IEEE Trans. Pattern Recogn. Machine Intell. **19**(4), 302–314 (1997)
6. Kovacs-Vajna, Z.M.: A fingerprint verification system based on triangular matching and dynamic time warping. IEEE Trans. Pattern Recogn. Machine Intell. **22**(11), 1266–1276 (2000)
7. Chen, X.J., Tian, J., Yang, X.: A new algorithm for distorted fingerprints matching based on normalized fuzzy similarity measure. IEEE Trans. Image Process. **15**(3), 767–776 (2006)
8. Bazen, A.M., Gerez, S.H.: Fingerprint matching by thin-plate spline modelling of elastic deformations. Pattern Recogn. **36**(8), 1859–1867 (2003)
9. Ross, A., Dass, S., Jain, A.K.: A deformable model for fingerprint matching. Pattern Recogn. **38**(1), 95–103 (2005)
10. He, Y.L., Tian, J., Li, L., Yang, X.: Fingerprint matching based on global comprehensive similarity. IEEE Trans. Pattern Analy. Machine Intell. **28**(6), 850–862 (2006)
11. Jain, A.K., Chen, Y., Demirkus, M.: Pores and ridges: high-resolution fingerprint matching using level 3 features. IEEE Trans. Pattern Recogn. Machine Intell. **29**(1), 15–27 (2007)
12. Jain, A.K., Prabhakar, S., Lin, H., Pankanti, S.: Filterbank-based fingerprint matching. IEEE Trans. Image Process. **9**(5), 846–859 (2000)
13. Gu, J., Zhou, J., Yang, C.: Fingerprint recognition by combining global structure and local cues. IEEE Trans. Image Process **15**(7), 1952–1964 (2006)
14. Wang, X.C., Li, J.W., Niu, Y.M.: Fingerprint matching using OrientationCodes and PolyLines. Pattern Recogn. **40**(11), 3164–3177 (2007)
15. http://bias.csr.unibo.it/fvc2002/perfeval.asp
16. Cappelli, R., Maio, D., Maltoni, D., Wayman, J.L., Jain, A.K.: Performance evaluation of fingerprint verification systems. IEEE Trans. Pattern Recogn. Machine Intell. **28**(1), 3–18 (2006)
17. http://fpvte.nist.gov
18. Ross, A., Jain, A.K.: Biometric sensor interoperability: a case study in fingerprints. Proc. Biometric Authentication: ECCV 2004 International Workshop **3087**, 134–145 (2004)
19. Bolle, R.M., Ratha, N.K., Pankanti, S.: An evaluation of error confidence interval estimation methods. Fifteenth ICPR **3**, 103–106 (2004)
20. Ross, A., Jain, A.K., Reisman, J.: A hybrid fingerprint matcher. Pattern Recogn. **36**(7), 1661–1673 (2003)
21. Bhakar, S., Jain, A.K.: Decision-level fusion in fingerprint verification. Pattern Recogn. **35**(4), 861–874 (2002)
22. Chen, X.J., Tian, J., Yang, X. An algorithm for distorted fingerprint matching based on local triangle features set. IEEE Trans. Inform. Forensics Security **1**(2), 169–177 (2006)

Fingerprint Matching, Manual

Herman Bergman[1], Arie Zeelenberg[2]
[1]Certified Fingerprint Expert, San Francisco, USA
[2]Senior Advisor Fingerprints, National Police Force, The Netherlands

Synonyms

Identification; Individualization; Minutial

Definition

Identification has been defined as the determination by a fingerprint examiner that two examined images of friction ridge skin are deposited by the same source (finger, palm or foot), with the goal of determining the identity of a donor. If this can be established it is generally accepted within the discipline that given the uniqueness or ► individuality of friction ridge skin, this fingerprint can be attributed to this donor at the same time excluding all others. (In this contribution an expert for practical reasons is referred to as "he". Female experts should not feel excluded but may comfort themselves with the idea that with respect to erroneous identifications also the male form is used)

Fingerprint Matching: Manual

The matching process described here applies to marks or latent prints found at a crime scene or on pieces of evidence associated with a crime. Those marks tend to be incomplete and of lesser quality than ► comparison prints. The process where known prints are compared, one to one or one to many, to verify an identity has become an increasingly automated process. Because of the amount of quality and quantity of data available and the accuracy of current Automated Fingerprint Identification Systems (AFIS) this process can be applied in a "lights out" mode or monitored by examiners.

This automated process to determine individuality is generally referred to as "matching" and is executed by matching algorithms. For the process where latent prints or marks are analyzed and compared by an examiner the more generic term identification or individualization is used rather than matching.

The identification process is a one to one comparison and starts after a similar print is found which cannot be excluded as being the same at face value. Three possible scenarios can lead to this:

A candidate can be the result of an AFIS search in which the similarity of the extracted features is calculated against known exemplars in a digital repository.

If one of the best resembling candidates cannot be excluded it might be eligible for input in an identification process.

Second, a candidate can be selected after manual comparison of one or more named suspects.

Third, a candidate may be found through a manual search of a physical fingerprint repository. This last occasion becomes increasingly rare because physical fingerprint repositories and manual searching become distinct by the broad use of AFIS.

The process by which the expert examines possible candidates focuses more on elimination based on differences than on weighing of similarities. At this stage the examiner searches for differences in the overall pattern formed by the ridges which is considered the first of three levels of information that are generally distinguished [1]. They are addressed to as ▶ the first, ▶ second and ▶ third level detail.

When an expert manually compares a mark against known, or comparison prints he visually assesses the main aspects of the ridge flow and/or a discernible pattern and a chosen target group of ▶ minutiae which he can relate to a recognizable area or location in the mark such as a delta, core or along the type lines.

This information is used to eliminate compared prints, this exclusion may be a very fast process. At one glance an expert may see that a compared donor shows 10 whorl patterns in the fingertips while he is looking for a loop. Even so a donor with a number of loops to the right with high ridge counts between the delta and core can be excluded definitively if the mark has a low ridge count. If no exclusion on ridge flow is possible because it is similar the remaining print will be compared keeping the target group in mind and looking for differences in the known print at the given positions relative to known locations. If he initially finds small clusters in a similar sequence he will then expand the assessed area both in the mark and the known exemplar.

If the print does not originate from the same source he will quickly find discrepancies and the comparison print will be excluded. If exclusion fails, the candidate will be included in the identification process.

The identification process

The generally accepted methodology for the identification process of friction ridge impressions is known as ACE-V [1] or a variation of this [2]. ACE-V is the acronym for Analysis, Comparison and Evaluation followed by Verification by another expert. ACE-V was first introduced by R.A. Huber [3] and later by D. Ashbaugh [1] for the examination of friction ridge skin. This methodology is generally accepted in forensics as a universal protocol to promote reproducibility and objectivity and should allow for the validation of the stated conclusions by reference to the process through which they are constructed.

It has been argued that ACE-V may not fully provide the requirements [4] necessary for an identification technique which should be explicit and defined in more detail [5]. Professors van Koppen and Crombag [6, 7] proposed the use of a descriptive model and a decision making model in forensic identification of ear-, lip- and fingerprints.

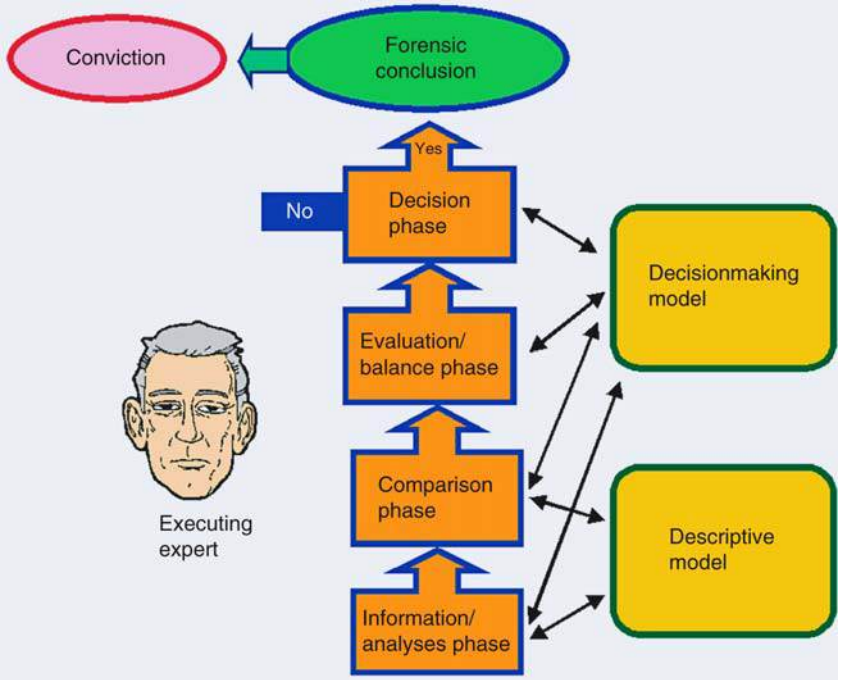
The Interpol European Expert Group on Fingerprint Identification (IEEGFI) report [2] not only describes a method similar to ACE-V (Fig. 1), but also provides both a descriptive model and a decision making model [6].

These models present a common terminology, grounds for establishing the value of features, rules of thumb, describe the pitfalls and provide good guidance for decision-making with respect to details and the overall decision of identification. It is essential to reproduce the whole process rather than to confine reproducibility to the conclusion.

The IEEGFI uses the word information phase as a synonym for the analyses phase and addresses the evaluation as the balance phase.

Analyses

A thorough and objective analysis of the latent print is the basis of a sound process, an unbiased establishment of the quantity and quality of available data is the aim. The analysis is the establishment of features and their properties and values recorded in a combination of mental and explicit written notes of all observed data. A copy of the image of the latent can be marked up in order to document observations. All three levels of information that are regarded as properties of friction ridge skin are assessed to determine their reliability



Fingerprint Matching, Manual. **Figure 1** Diagram ACE.

and value, taking into account the influence of development technique(s) used, the exhibit, distortion, surface, deposition pressure, matrix, and anatomical aspects. Ambiguous Galton features of which the exact location cannot be seen at face value can still be established by ▶ tracing. In these instances the ridge detail and the exact appearance of the detected feature are unknown and may add little weight to the value of the latent and, subsequently, to the comparison. Nevertheless, it can be helpful to check whether certain Galton features in the comparison print are at least not in conflict with the latent.

Although a good practise in all cases, it is acknowledged that not all latent prints require such an in-depth analysis. In instances of high quality latent prints with unambiguous and/or an abundance of data, the analysis can be very quick.

However, it should be stressed that with low quality and quantity latent prints a full in-depth analysis is essential. The importance and depth of the analysis is inversely proportional to the quality of the latent print.

The IEEGFI II proposed a special procedure, "▶ The need for a questionable ID procedure" for complex examinations [6]. The examiner has to form an opinion about the quality, quantity, and reliability of the observed data in the latent print and on the basis of this he has to decide whether the latent print has sufficient potential to relate it to its unique source. If that is the case he moves on to the comparison phase.

Comparison

The latent and the comparison prints are placed side by side enabling accurate comparison and the preservation of observations.

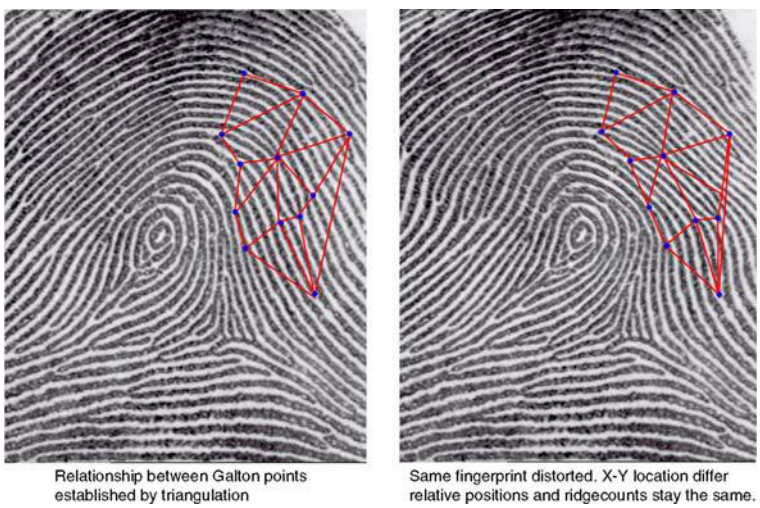
The data obtained in the analysis phase form the basis and guide for the comparison process and should be leading. During comparison not only data in the latent are checked against the comparison print, but also data found in the comparison print are cross-checked whether or not they are present in the latent.

The relations of all features within the configuration are checked through triangulation [8]. This is done by following the ridges or furrows and counting the number of intervening ridges between features along a virtual connecting line. The relative location aspects and relations to other features in the latent have to be within tolerance compared to the features in the corresponding locations of the comparison print. (The direction in which the neighboring feature is found is

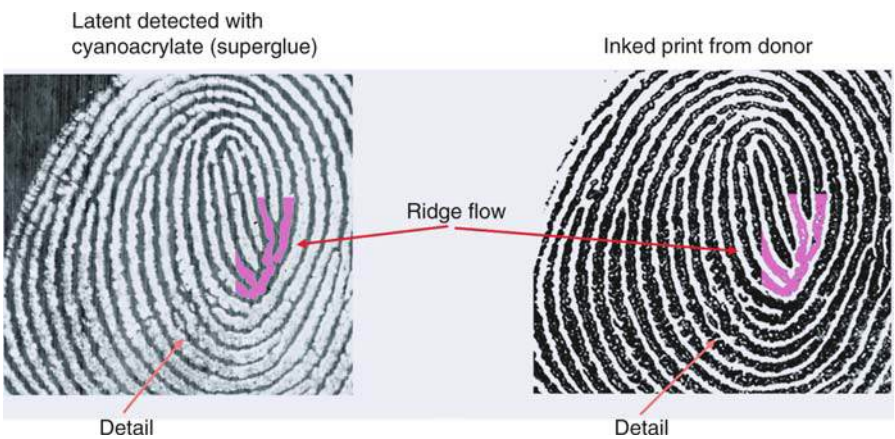
checked towards the general ridge flow and relative to the connecting line with other minutiae in the same area.) Due to the flexibility of the skin the interrelationship of features can be disturbed, but as in the case of a stretched spider web the relative positions remain the same (Fig. 2).

At this stage the expert also looks for similar third level detail which he relates to the location of second level detail.

It can be a very powerful contribution to each individual minutia and to the whole of a print but its accurate representation is dependent on a large number of variables such as pressure, moisture, the surface, and the detection technique. Reliable third level detail in latent prints is a gift rather than a given fact. It is often difficult to draw a distinction between third level detail and anomalies. Matching third level detail is not very common and often calls for rationalization. Instead, the relationship of the minute events amongst them and with minutiae is more often studied. A large pore on the edge of a line followed by a small one in the centre of a ridge, the flow of an individual ridge like a recognizable river bed, a small dot lying in front of a tapered ridge ending are examples of ridge detail that, if similar, can be very significant contributions to the weight of the comparison (Fig. 3).



Fingerprint Matching, Manual. **Figure 2** Triangulation/Distortion.



Fingerprint Matching, Manual. **Figure 3** Third level detail of exceptional quality, organic shapes that are found in a similar fashion on exact corresponding locations.

Overall similarities should be apparent and demonstrable and be primarily based upon findings obtained in the analysis phase. Thus, avoiding the implementation of features found in the supposed original into the latent. When marking corresponding features it is important to establish the existence, the relations and their significance. For each individual point of similarity the quality may differ. If a point is clear and shows corresponding ridge detail its value is significantly higher than points that do not have these properties.

Dissimilarities and/or discrepancies should be detected, assessed, noted and accounted for. Any explanation of dissimilarities should preferably reflect the observations made during the analysis phase. An opinion has to be formed whether the differences in appearance are considered distortions or discrepancies. In the case of discrepancies, the conclusion should be an exclusion or/and inconclusive.

There is a distinct difference between the comparison of minutiae and ridge detail. Minutiae must be the same and ridge detail can be the same. Whereas the basic properties of the Galton points are firmly established during the analysis phase true third level detail is often only acknowledged and confirmed during comparison taking the supposed original as the blueprint. This carries the risk of a picking attitude of the expert who may select everything that appears to be similar and ignore all that is not.

Further, this promotes the risk of circular reasoning [6] or "gestalt analyses" [4], instead of proving origin by the similarities one "proves" similarities by the assumed origin.

It has been discussed that the ACE-V protocol is a recurring and reversible process [9, 10]. Opinions vary however whether or not the process should be totally recurring, and reversible (or up to a certain level) or that attempts should be made to confine it to a more linear process wherever possible. With a recurring and reversible process the risk of inserting information of the "known" exemplar into the unknown is higher than in a strict linear process in which ACE-V is executed once in the exact order.

The risk of making a (subconscious) decision early in the comparative process and the potential influence of it must be recognized [11]. The comparison must be an unbiased "step by step" building process ensuring that the data in the latent and comparison print match, with nothing in disagreement which cannot be logically explained and accounted for. The decision must be made at the end of the process only.

An expert who has executed the process of searching and elimination has performed an initial and incomplete analysis directed towards elimination and/or the search process. Since he has singled out a comparison print for the identification process he has arrived at a preliminary conclusion about possible identification. With an eye to the "half baked" analysis and the preliminary conclusion it is advisable that the expert renounces himself from the identification process.

Evaluation and Preliminary Conclusion

Requirements for the conclusion of identification as provided by SWGFAST [12] are; agreement of sufficient friction ridge detail; determined by a competent examiner; applied to a common area in both impressions; based on quantity and quality of friction ridge detail; without any discrepancy and a reproducible conclusion. The total volume in agreement is a composition of coherent qualitative and quantitative information.

In the USA, after a 3-year study by a Standardization Committee, the use of a numerical standard was discouraged by the adoption of a resolution at a conference of the International Association for Identification which stated: "no scientific basis exists for requiring that a predetermined minimum of friction ridge features must be present in two impressions in order to establish positive identification".

Sufficiency is since left to the discretion of the expert and measured against his training, knowledge and experience, and his personal standard. SWGFAST [12] relates reproducibility primarily to the format of ACE-V and to the conclusion. This position is known as the Expert Opinion System or the holistic approach [2, 13].

In many other countries a numerical standard is used as an aid to measure sufficiency which is called the Empirical Standard Approach [2, 13]. This standard expresses a minimum number of minutiae in agreement that is used as a common, empirical reference and a tool to guide the process, to facilitate verification and to obtain and guarantee quality.

In either system if an expert decides that in his opinion identification is justifiable because the equation is both sufficient and cogent he will put it up for verification [14].

Verification

The postulated conclusion should be reproducible by another examiner applying the same methodology. This is accomplished by the verification phase of the ACE-V methodology.

The reliability of a conclusion can be checked and demonstrated by an independent verification. Verification can be limited to another expert independently arriving at the same conclusion or by repeating and checking the whole examination of the initial examiner. The verification process should have the characteristics of scrutiny rather than confirmation of the conclusion. (Also see mistakes.)

If the verifier is satisfied that the process and the conclusion meet the requirements, then the conclusion is confirmed and the identification is established.

Conclusion

The conclusion of identification is a verified opinion that the investigated latent and the comparison print come from the same source. It also implicates the expectation of reproducibility, i.e., any other examiner using the same methodology should arrive at the same conclusion. Given the empirical, biological and statistical support for friction ridge skin uniqueness or individuality, an identified fingerprint is attributed to a single donor [15].

Charting

The use of a computer screen during the analysis and side by side comparison of friction ridge images can be of tremendous help in the examination process. The data and the relations of the configuration can be cross-checked, in particular with ambiguous information. Details can be better observed and compared by enlarging and/or enhancement of the images to optimize the perceptibility of the characteristics in print. This not only increases the quantity and accuracy of the data observed [16], but also makes it easier to value and appreciate the similarities and dissimilarities. At the same time similarities can be marked up, printed and saved for documentation purposes.

In order to meet the requirement of demonstrability of all the phenomena upon which the expert bases his findings and conclusion this tool is indispensable. It also facilitates consultation and discussion amongst experts.

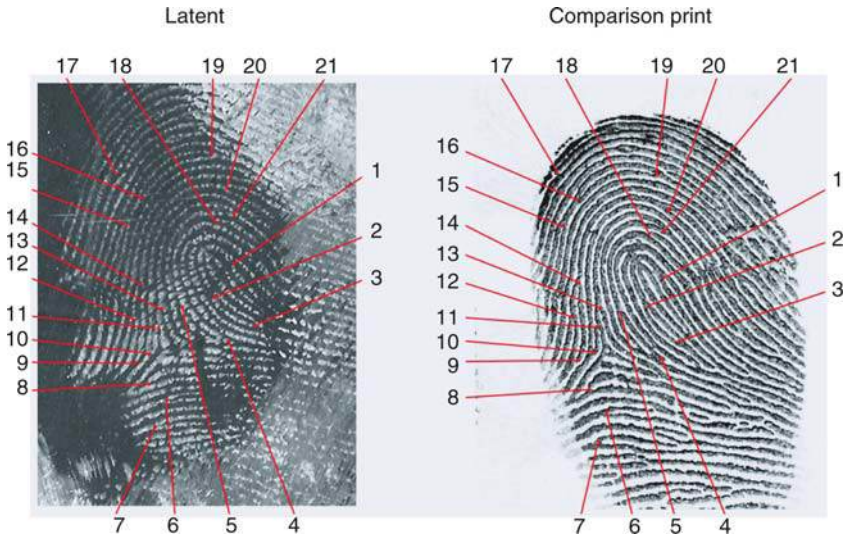
Historically, court charts have been produced in which coinciding minutiae have been marked up and numbered. Court charts can be a useful tool to demonstrate some of the findings but are just meager illustrations of a very complex process and should not be taken as ultimate proof. (The simple argument for that is the fact that in the past with erroneous identifications court charts were produced with even extensive numbers of marked similarities (Fig. 4).)

Mistakes

Error rates, an endless source of scientific debates, philosophies and semantics, will not be covered nor decided here but known errors will be discussed. It is obvious that, in relation to the immense numbers of identifications effected over the more than hundred years of fingerprinting, the number of erroneous identifications that surfaced is extremely low. In a study by Simon Cole [17] 22 erroneous identifications were investigated for the period from 1920 to 2004. Even if the number is tripled or multiplied by 10 incorporating a number of dark figures the positive ratio against the millions of identifications performed remains. Some support for this positive ratio is also found in data collected during comparison training exercises [18].

This does not implicate that mistakes are regarded as part of the system and inevitable, on the contrary. Every mistake is one too many and can do irreparable harm to innocent people. The profession should take all possible measures to prevent them.

Another ongoing debate is whether the mistakes can be attributed to flaws in the technique or to human error or whether the two can be separated at all. However, it is clear that erroneous identifications are discovered and exposed by experts. This is a strong indication that the human factor is dominant. Second, it is important to note that the examination of questioned identifications can be repeated and checked endlessly. When opinions differ upon sufficiency the comparsion should be regarded as inconclusive. In general, experts view mistakes very seriously. They believe that making a mistake is the worst thing that can happen to them and may discredit them in the eyes of their contemporaries.



Fingerprint Matching, Manual. **Figure 4** Example of a court chart with 21 coinciding Galton characteristics marked.

The paradox is that the acceptance of the susceptibility for human error by experts should be the basis for a quality system, whilst very often the initial response of experts to such criticism is defensive rather than open minded. This blocks the feedback essential for a quality cycle.

A preliminary analysis of mistaken identifications revealed the following factors:

- It concerned a border line latent with respect to quality, quantity or both.
- There was no apparent relationship between the organization or level of experience of the expert (s) involved.
- Verification was degenerated to confirmation rather than scrutiny.
- Experts were biased by domain irrelevant information.
- Discrepancies were ignored or erroneously attributed to distortion.
- Applied tolerances were too wide given the quality of the latent. This is another paradox; "the worse the print the larger the tolerances applied"; experts may attribute differences to the lack of quality and distortion and "explain them away" something they may not do with an image of good quality. Thus, bad quality may not only conceal real discrepancies, but also provide an excuse for it at the same time.

In general, there is a growing opinion that a number of psychological factors may potentially contribute to cognitive and decision-making errors [19].

Examples are; the primacy effect, when information is judged in the light of an early opinion; and confirmation biases like myside bias and truth bias [11] are found in all types of fields as well as in ordinary life. One major concern is that sufficiency may be established after the comparison process and as such after a conscious or subconscious decision is made about identity. This makes the expert more vulnerable to bias [19].

Studies have been done to enhance insight into the potential influences of bias during the examination of fingerprints [20].

Infallible or Reliable?

Some have criticized the profession for the explicit or implicit claim of infallibility [17, 21].

The apparent reliability of fingerprint identification for decades may have created this image as reflected by the proverbial expression "as reliable as a fingerprint". This meant an image so strong that all other forensic techniques were compared against it, much like the introduction of DNA that was erroneously labeled the "genetic fingerprint".

Responsible experts never claim infallibility because this is an unsustainable and unscientific position.

In retrospect, however, fingerprints in general can claim a record of great reliability, but as in any human endeavor mistakes occur so safeguards have to be in place.

The main ground for quality is the acceptance of fallibility by individuals and communities. With that in mind, instruments to achieve a solid conclusion, the rigorous application of the methodology, Quality Assurance protocols, training, testing and transparency, will be applied and maintained with conviction and can be further improved.

Per individual case reliability of a conclusion can be reached and demonstrated by verification, peer review and counterchecks by independent experts. This process can be repeated over and over again without affecting the material.

Related Entries

- ▶ Classification
- ▶ Feature Extraction
- ▶ Fingerprint Classification
- ▶ Fingerprint Matching Automatic
- ▶ Individuality

References

1. Ashbaugh, D.: Quantitative-Qualitative friction ridge analysis, p. 105. CRC Press, Boca Raton, FL (1999)
2. Method for Fingerprint Identification, Interpol European Expert Group for Fingerprint Identification Report I http://www.interpol.int/Public/Forensic/fingerprints/WorkingParties/default.asp
3. Huber, R.A.: Expert witness. Criminal Law Quarterly. **2**, 276–295 (1959)
4. Rudin, N., Inman, K.: The proceedings of lunch The CAC News of the California Association of Criminalists, 2nd Quarter (2005) http://www.cacnews.org/
5. A Review of the FBI's Handling of the Brandon Mayfield Case, Office of the Inspector General, http://www.usdoj.gov/oig/special/s0601/PDF_list.htm pp. 7, 198–199. (2006)
6. Method for Fingerprint Identification, Interpol European Expert Group for Fingerprint Identification Report II. http://www.interpol.int/Public/Forensic/fingerprints/WorkingParties/default.asp
7. van Koppen, P.J., Crombag, H.H.: Over Oren, Lippen en Vingers. Nederlands Juristenblad (2000)
8. Hare, K.: Proportional analysis: The science of comparison. J. Forensic Ident. **53**, 700 (2003)
9. Vanderkolk, J.R.: ACE-V: A Model. J Forensic Ident. **54** (2004)
10. Mc Kasson, S.C., Richards, C.A.: Speaking as an expert: A guide for the identification sciences from the laboratory to the courtroom, 131–138 (1998)
11. Nickerson, Raymond S.: Confirmation bias: A ubiquitous phenomenon in many guises. Rev. Gen. Psychol. **2**, 175–220 (1998)
12. Scientific Working Group on Friction Ridge Analyses, Study and Technology http://www.swgfast.org/Standards_for_Conclusions_ver_1_0.pdf
13. C. Champod, et al., Fingerprints and other ridge skin impressions, 27–31 (2004)
14. Thornton, J.: "Setting Standards in the Comparison and Identification" In: 84th Annual Training Conference of the California State Division of IAI Laughlin, Nevada, May 9 (2000) http://www.latent-prints.com/Thornton.htm
15. Moenssens, A.A.: Is fingerprint identification a science? http://forensic-evidence.com/site/ID/ID00004_2.html
16. Langenburg, G.M.: A statistical analysis of the ACE-V methodology: Analysis stage. J. Forensic Ident. **54** (2004)
17. Cole, S.A.: More than zero: Accounting for error in latent fingerprint identification. J. Crim. Law Criminol. **95** (2005)
18. Wertheim, K., Langenburg, G. Moenssens, A.: "A report of latent print examiner accuracy during comparison training exercises". J. Forensic Ident. **56**, 55–93 (2006)
19. Itiel, D., Charlton, D., Péron, A.E.: "Why are experts prone to error?" Forensic Sci. Int. **156**, 74–78 (2006)
20. Schiffer B., Champod, C.: The potential (Negative) influence of observational biases at the analysis stage of fingermark individualization. Forensic Sci. Int. **167**, 116–120 (2007)
21. Saks, M., Koehler, J.: The coming paradigm shift in forensic identification science. Science, **309**, 892 (2005)

Fingerprint Pre-Matching

▶ Fingerprint Classification

Fingerprint Quality

The intrinsic characteristic of a fingerprint image that may be used to determine its suitability for further processing by the biometric system or assess its conformance to pre-established standards is fingerprint quality. The quality of a biometric signal is a numerical value (or a vector) that measures this intrinsic attribute.

▶ Individuality of Fingerprints

Fingerprint Reading

► Biometric Sample Acquisition

Fingerprint Recognition, Overview

Davide Maltoni
DEIS, University of Bologna, Italy

Synonym

Fingerprint Biometric

Definition

Fingerprint recognition allows a person to be verified or identified through the analysis and comparison of his or her finger dermal ridges. Fingerprint recognition was one of the first techniques used for automatically identifying people and today is still one of the most popular and effective biometric techniques.

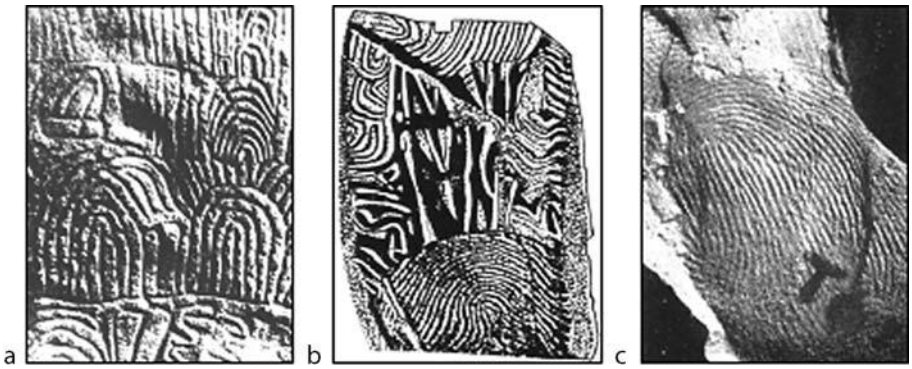
Introduction

A fingerprint is the representation of the dermal ridges of a finger [1]. Dermal ridges form through a combination of genetic and environmental factors; the genetic code in DNA gives general instructions on the way skin should form in a developing fetus, but the specific way it forms is a result of random events such as the exact position of the fetus in the womb at a particular moment. This is the reason why even the fingerprints of identical twins are different [2]. Fingerprints are fully formed at about 7 months of fetus development and finger ridge configurations do not change throughout the life of an individual, except in case of accidents such as severe cuts on the fingertips. This stability makes fingerprints a very attractive biometric identifier. Several mathematical models based on the ► anatomy of friction ridge skin were developed over the years to quantify ► fingerprint individuality [3] and to prove that finding two persons with identical fingerprints is extremely unlikely. This does not imply that fingerprint recognition is a perfect technique: in fact, various kinds of errors can affect fingerprint acquisition and processing thus requiring to introduce thresholds to decide if two fingerprint impressions are similar enough to be considered belonging to the same person. As for any biometric technique, a sound performance evaluation (see ► fingerprint databases and evaluation) is extremely important to estimate the accuracy of a fingerprint-based biometric system and to understand if it is well-suited for a particular application. Recent independent evaluation campaigns such as FVC2006 [4] proved that state-of-the art fingerprint recognition algorithms are nowadays very accurate (i.e., EER less than 0.1% for a database collected with a large area optical scanner).

History

Human fingerprints have been discovered on archaeological artefacts and historical items (Fig. 1). Although these findings prove that ancient people used fingerprints for a number of purposes, it was not until the late sixteenth century that the modern scientific fingerprint studies were initiated [5]. In 1686, Marcello Malpighi, a professor of anatomy at the University of Bologna, noted the presence of ridges, spirals and loops in fingerprints. Henry Fauld, in 1880, was the first to scientifically suggest the individuality of fingerprints based on an empirical observation. At the same time, Herschel asserted that he had practiced fingerprint recognition for about 20 years. In the late nineteenth century, Sir Francis Galton conducted an extensive study on fingerprints; in 1888 he introduced the ► minutiae features for fingerprint matching. Another important advance was made in 1899 by Edward Henry, who established the well-known "Henry system" of ► fingerprint classification.

In the early twentieth century, fingerprint recognition was formally accepted as a valid identification method and became a standard routine in forensics [5]. Fingerprint identification agencies were set up worldwide and criminal fingerprint databases were established; for instance, the FBI fingerprint identification division was set up, in 1924, with a database of 810,000 fingerprint cards. With the rapid expansion of fingerprint recognition in forensics, operational



Fingerprint Recognition, Overview. Figure 1 Examples of archaeological fingerprint carvings and historic fingerprint impressions: (**a**) Neolithic carvings (Gavrinis Island); (**b**) standing stone (Goat Island, 2000 B.C.); (**c**) an impression on a Palestinian lamp (400 A.D.). Figures courtesy of A. Moenssens and R. Gaensslen.

fingerprint databases grew so large that manual fingerprint identification (see ► fingerprint matching, manual) became infeasible; for example, the total number of fingerprint cards in the FBI fingerprint database stands well over 200 million and is continuously growing. With thousands of requests being received daily, even a team of more than 1300 fingerprint experts were not able to provide timely responses to these requests. Starting in the early 1960s, the FBI, Home Office in the UK, and Paris Police Department began to invest a large amount of effort in developing Automatic Fingerprint Identification Systems (► AFIS). Based on the observations of how human fingerprint experts perform fingerprint recognition, three major problems in designing AFIS were identified and investigated: digital fingerprint acquisition, local ridge feature extraction, and ridge characteristic pattern matching. Their efforts were so successful that today almost every law enforcement agency worldwide uses an AFIS. These systems have greatly improved the operational productivity of law enforcement agencies and reduced the cost of hiring and training human fingerprint experts.

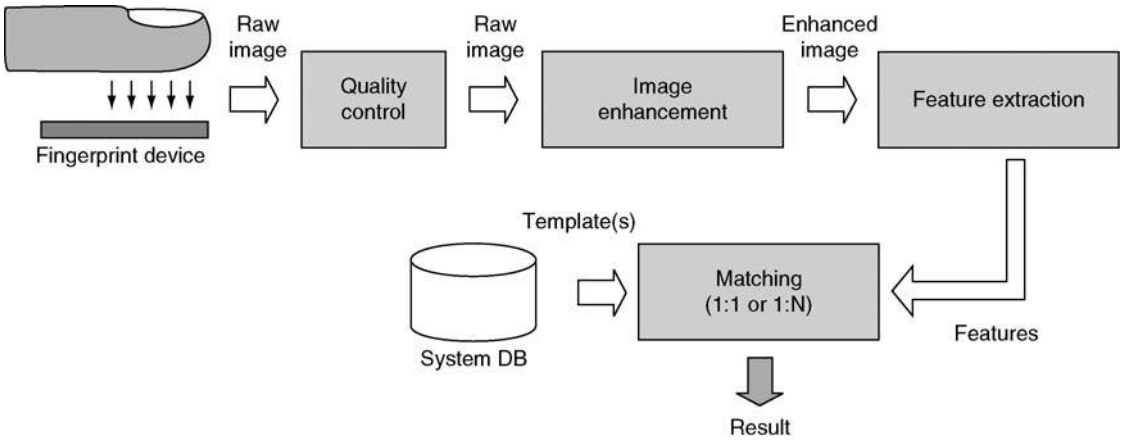
Automatic fingerprint recognition technology has now rapidly grown beyond forensic applications. On the one side, together with face, fingerprint is the main biometric modality for electronic documents (e-passport, visas, ID cards, etc) used to enforce border crossing and citizen security. On the other side, thanks to a very good performance/cost tradeoff, fingerprint-based biometric systems are becoming very popular and are being deployed in a wide range of commercial applications such as logon to computers and networks, physical access control, ATMs.

Components of a Fingerprint Recognition System

The block diagram of a fingerprint-based recognition system is depicted in Fig. 2.

A fingerprint is acquired through a live-scan ► fingerprint device that allows to simply and quickly capture a digital fingerprint image: most of the fingerprint devices sample the pattern at 500 DPI (Dots per Inch) and produce an 8-bit gray-scale raw image (see Fig. 3). Some devices also include fake detection mechanisms (see ► fingerprint fake detection) that allow to reveal spoofing attacks carried out with fake fingers.

The acquired raw image is then passed to a quality control module that evaluates if the fingerprint sample quality is good enough to correctly process it and to extract reliable features. In case of insufficient quality, the system rejects the sample and invites the user to repeat the acquisition; otherwise, the raw image is passed to an ► image enhancement module whose goal is improving the clarity of the ridge pattern, especially in noisy region, to simplify the subsequent feature extraction. Special digital filtering techniques, known as contextual filtering [1], are usually adopted at this stage; the output enhanced image can still be a gray-scale image or become a black-and-white image. The ► feature extraction module further processes the enhanced image and extracts a set of features from it. This feature set often includes minutiae but, depending on the matching algorithm, other features (e.g., local orientation, local frequency, singularities, ridge shapes, ridge counts, parts of the enhanced image, etc.) can be extracted in conjunction with (or instead of) minutiae.



Fingerprint Recognition, Overview. **Figure 2** Block diagram of a fingerprint-based recognition system.

Fingerprint Recognition, Overview. **Figure 3** Example of fingerprint images from FVC2006 databases [4].

Finally, the fingerprint matching module (see ► fingerprint matching, automatic) retrieves from a system database one or more templates (see ► fingerprint templates) and matches it/them with the features extracted from the current sample. Most of the matching algorithms, following the well established manual method (see ► fingerprint matching, manual), compare two fingerprints by searching the spatial correspondence of a minimum number of minutiae; this is not a simple task because of the large variations (e.g., displacement, rotation, skin condition, distortion, noise, etc.) that can characterize two fingerprint images acquired from the same finger at different times. If the systems is operating in verification mode, the user has been required to claim his identity and therefore just one template is retrieved from the database and matched with the current sample; if the system is operating in identification mode the current sample is matched against all the database templates to check is one of them is sufficiently similar.

Protecting fingerprint templates is very important to avoid attacks to fingerprint-based biometric systems [6] and to preserve user privacy: cryptography techniques can be used to this purpose (see ► Fingerprints Hashing).

Large-Scale Automatic Fingerprint Identification Systems

Large-scale automatic fingerprint identification systems (AFIS) are used in forensic and civil government applications. The basic functioning of these systems is the same as described in the previous section, but a number of ad-hoc optimizations are employed to effectively and efficiently store, retrieve and match millions of fingerprints in a few seconds. In the past, special dedicated hardware and storage devices were used to guarantee the required throughput; nowadays, most of the AFIS cores run on conventional hardware (e.g., cluster of personal computers) and the software is the main responsible of the system efficiency. Fingerprint classification and ► fingerprint indexing are the two main techniques used to speed-up a fingerprint search in a large database [1]. The former allows to split the database in a number of partitions and to limit the search to the partition to which the searched sample belongs to. The latter enables sorting the database templates according to the similarity with the searched

sample, so that the probability to find a mate in the first attempts increases significantly. Even if the capacity of mass storage devices is continuously growing, storing fingerprints as uncompressed raw images would require too much space (nowadays AFIS must store billions of fingerprint images) and would increase the time necessary to transmit a fingerprint record over a network; to alleviate this problem, without compromising recognition accuracy, specific ▶ fingerprint compression techniques such as WSQ (Wavelet Scalar Quantization) have been developed by researchers.

Related Entries

- ▶ Biometrics, Overview
- ▶ Biometric Recognition
- ▶ Fingerprint Anatomy
- ▶ Fingerprint Classification
- ▶ Fingerprint Compression
- ▶ Fingerprint Databases and Evaluation
- ▶ Fingerprint Fake detection
- ▶ Fingerprint Features
- ▶ Fingerprint Image Enhancement
- ▶ Fingerprint Indexing
- ▶ Fingerprint Individuality
- ▶ Fingerprint Matching, Automatic
- ▶ Fingerprint Matching, Manual
- ▶ Fingerprint Image Quality
- ▶ Fingerprint Templates
- ▶ Fingerprints Hashing

References

1. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of fingerprint recognition. Springer, New York (2003)
2. Jain, A.K., Prabhakar, S., Pankanti, S.: On the similarity of identical twin fingerprints. Pattern Recognit. **35**(11), 2653–2663 (2002)
3. Pankanti, S., Prabhakar, S., Jain, A.K.: On the individuality of fingerprints. IEEE Trans. Pattern Anal. Mach. Intell. **24**(8), 1010–1025 (2002)
4. The Fourth International Fingerprint Verification Competition (FVC2006) http://bias.csr.unibo.it/fvc2006.
5. Lee, H.C., Gaensslen, R.E.: Advances in fingerprint technology. 2nd Edn. Elsevier, New York (2001).
6. Cappelli, R., Lumini, A., Maio, D., Maltoni, D.: Fingerprint image reconstruction from standard templates. IEEE Trans. Pattern Anal. Mach. Intell. **29**(9), 1489–1503 (2007)

Fingerprint Representation

▶ Fingerprint Templates

Fingerprint Retrieval

Fingerprint retrieval is a procedure that draws a subset of fingerprints from a database stored on a computer system based on some similarity measure between a query fingerprint and the fingerprints in the database. The ultimate goal of fingerprint retrieval is not to find a group of fingerprints similar to the query fingerprint, but to get back the fingerprint originating from the same finger as that of the query fingerprint. Hence, success or failure of the fingerprint retrieval is determined by whether the retrieved subset contains the fingerprint originating from the same finger as that of the query fingerprint.

- ▶ Fingerprint Classification
- ▶ Fingerprint Indexing

Fingerprint Sample Synthesis

Raffaele Cappelli
Biometric System Laboratory, DEIS, University of Bologna, Cesena, Italy

Synonyms

Synthetic fingerprint generation; Synthetic fingerprints; Artificial fingerprints

Definition

Fingerprint sample synthesis is the generation of images similar to human fingerprints, through parametric models that simulate the main characteristics of such biometric data and their modes of variation. The image synthesis is typically performed by a computer program that, starting from some input

parameters, executes a sequence of algorithmic steps that finally produce a synthetic fingerprint image.

Introduction

With the increasingly adoption of fingerprint recognition systems, driven by their very appealing accuracy/cost tradeoff, methodical and accurate performance evaluations of fingerprint recognition algorithms are needed. Unfortunately, this requires large databases of fingerprints, due to the very small rates of error necessary for the procedure. For instance, according to [1], in order to support a claim of FMR less than 1/10,000 (the requirement for verification applications in [2]), 30,000 impostor matches from at least 250 individuals should be performed without observing any false match error. On the other hand, collecting large databases of fingerprint images is expensive both in terms of money and time, boring for both the people involved and for the volunteers, and problematic due to the privacy legislation that protects such personal data. FVC competitions [3] are examples of technology evaluations, where real fingerprint databases have been collected to test different algorithms, but do not constitute lasting solutions for evaluating and comparing different algorithms; in fact, since FVC databases are made available to the participants after the competition to let them improve the technology, they expire once "used," and new databases have to be collected for future evaluations.

Fingerprint synthesis is a feasible way to address the issues just cited, since it allows large databases of images to be easily generated and used for testing fingerprint recognition systems without infringing on privacy.

A fingerprint synthesis method typically consists of two main steps: first, a ridge pattern, which represents the unique and immutable characteristics of a "synthetic finger," is generated according to a given model; then, one or more "fingerprints" of the synthetic finger are generated by simulating the main factors that make the fingerprints of a given human finger different each other.

Physical Ridge Pattern Models

Physical ridge pattern models are based on some hypothesized physical mechanisms of fingerprint formation during embryogenesis.

The crucial period of fingerprint development in humans starts at the 10th week of pregnancy [4], when the epidermis consists of three layers (outside layer, intermediate layer and basal layer). It is then observed that the basal layer of the epidermis becomes undulated toward the surface, forming the so-called "primary ridges," whose development ends at about the 17th week of pregnancy: at this stage the geometry of the epidermal ridge pattern is determined for life and becomes visible on the skin surface in subsequent weeks.

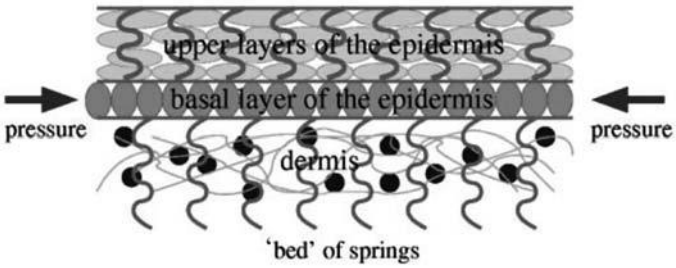
Several theories for fingerprint pattern formation have been proposed in the scientific literature [4], including cell proliferation phenomena, mechanical interaction between the extracellular matrix and fibroblasts in the dermis, reaction-diffusion models.

In a study by Sherstinsky and Picard [5], a complex method which employs a dynamic non-linear system called "M-lattice," is introduced. The method is based on the reaction-diffusion model first proposed by Turing in 1952 to explain the formation of animal patterns such as zebra stripes. Although this work is aimed at optimally binarizing a fingerprint image, the underlying ridge-line model could be used as a basis for synthetic generation.

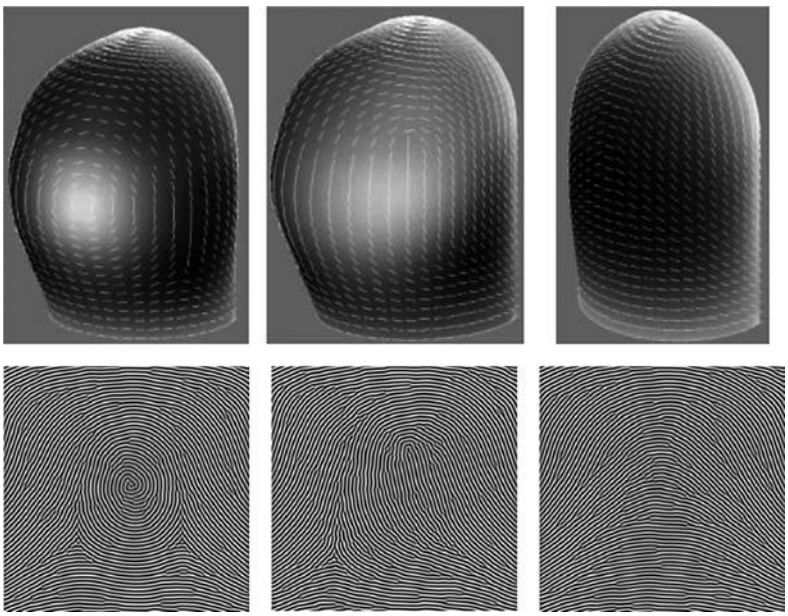
An interesting model was proposed by Kücken [4, 6], based on the following hypotheses:

1. Fingerprint patterns are created by forces that are induced by differential growth of the epidermis' basal layer (as argued by Cummins [7] from the observed dependency of the pattern class on the fingertip geometry)
2. Non-uniform growth of the epidermis' basal layer results in compressive stress that leads to buckling, creating the primary-ridges [8]

Kücken considers the basal layer as an elastic sheet trapped between the neighboring tissues of the intermediate epidermis layer and the dermis (Fig. 1) and studied the buckling process by means of the von Karman equations, which describe the behavior of a thin curved sheet of elastic material. The analysis of those equations confirmed that the direction of the ridges is roughly perpendicular to the direction of greatest stress; Kücken postulated that two factors mainly contribute to generate the compressive stress in the basal layer: (1) resistance at the nail furrow and at the major flexion creases of the finger (boundary effects); (2) the regression of the "volar pads" at the time of fingerprint development. Volar pads are



Fingerprint Sample Synthesis. **Figure 1** The basal layer of epidermis: Kücken and Newell [6] assumes that due to differential growth, a compressive stress act on this layer.



Fingerprint Sample Synthesis. **Figure 2** Simulation of three common fingerprint patters (from left to right: whorl, loop, and arch) using the model proposed in [6].

temporary eminences of the skin surface that form during the 7th week of pregnancy and start to digress at about the 10th week. From studies of embryos, monkeys and malformed hands, it has consistently been observed that highly rounded pads at the fingertips exhibit whorls; less well-developed pads show loops, where the direction of the loop opening is determined by the asymmetry of the pad; small indistinct pads give rise to arches.

Computer simulations have shown results consistent with the above observations and hypothesis; Fig. 2 shows how an almost periodic pattern very similar to human fingerprints can be generated by applying Kücken's model: the three main fingerprint classes can be simulated and ▶ minutiae are present in regions where ridge patches with different directions and/or wavelength meet.

Statistical Ridge Pattern Models

Statistical ridge pattern models aims to reproduce realistic-looking fingerprints without starting from embryological hypothesis. Such models are based on the empirical analysis of real fingerprints, from which statistical data about the main characteristics of the patterns are derived and parameterized into appropriate equations or synthesis algorithms.

In 1999, Kosz published some interesting results concerning fingerprint synthesis based on a

mathematical model of ridge patterns and minutiae [9]; further details on this technique have been provided online by Bicz [10] in 2003. According to this model, a fingerprint can be described by a wave pattern:

$$f(x, y) = \cos(\varphi(x, y)) \tag{1}$$

where:

$$\varphi(x, y) = \varphi_0(x, y) + \varphi_M(x, y) \tag{2}$$

is a function that defines the phase of the wave structure as the sum of two parts: φ_0 , which describes the global "shape" of the ridge lines, and φ_M, which describes the minutiae. According to the model introduced by bicz [10], φ_M can simply generate n minutiae by adding n spatially-shifted arctangent functions:

$$\varphi_M(x, y) = \sum_{i=1}^{n} \arctan\left(\frac{y - y_i}{x - x_i}\right) \tag{3}$$

where (x_i, y_i) is the location of minutia i. Figure 3 shows a synthetic pattern generated by using the above equations.

In 1993, Sherlock and Monro [11] proposed an orientation model that allows a consistent ▶ orientation field to be computed from the sole knowledge of the position of fingerprint ▶ singularities (loops and deltas). In this model, the image is located in the complex plane and the local ridge orientation is the phase of the square root of a complex rational function whose singularities (poles and zeros) are located at the same place as the fingerprint singularities (loops and deltas). Let $\mathbf{ls}_i$, $i = 1..n_c$ and $\mathbf{ds}_i$, $i = 1..n_d$ be the coordinates of the loops and deltas respectively. The orientation θ at each point $\mathbf{z} = [x, y]$ is calculated as:

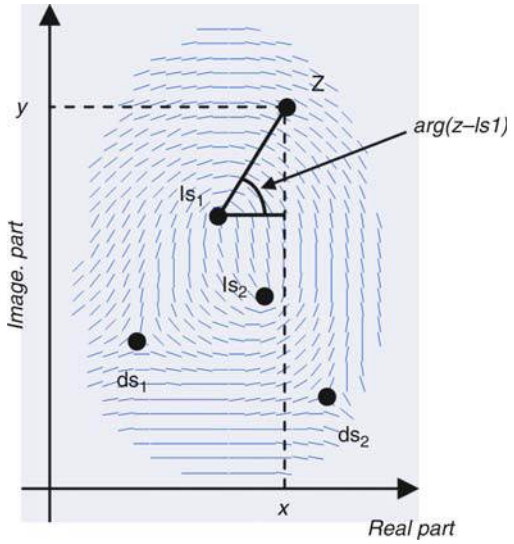
$$\theta = \frac{1}{2}\left[\sum_{i=1}^{n_d} arg(\mathbf{z} - \mathbf{ds}_i) - \sum_{i=1}^{n_c} arg(\mathbf{z} - \mathbf{ls}_i)\right] \tag{4}$$

where the function $arg(\mathbf{c})$ returns the phase angle of the complex number $\mathbf{c}$ (see Fig. 4).

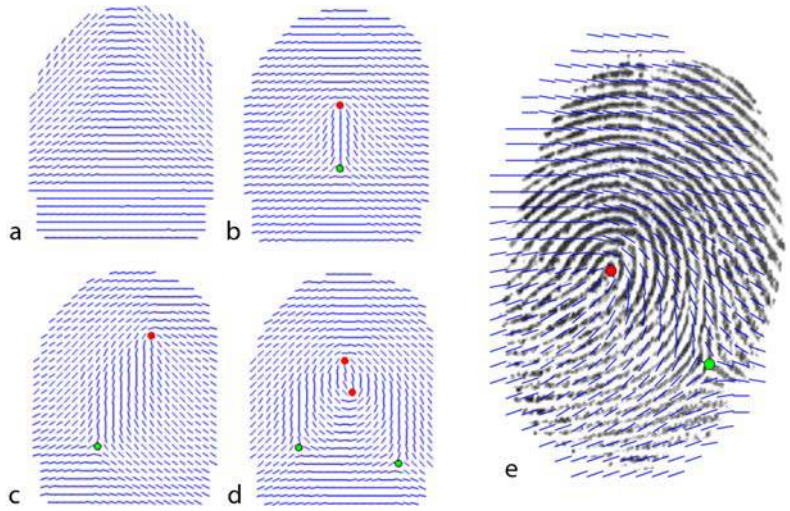
The Sherlock and Monro model may be exploited for generating synthetic orientation fields by first randomly choosing a fingerprint class and then randomly selecting the positions of the singularities, according to the class-specific constraints (for instance, in a left loop, the delta must be on the right side of the loop). Figure 5 shows some examples of orientation fields generated by this model.

However, in nature the ridge-line flow cannot be completely determined by the singularity type and position. In 1996, Vizcaya and Gerhardt proposed a variant of the Sherlock and Monroe model that

Fingerprint Sample Synthesis. Figure 3 A simple synthetic pattern generated by equations (1)–(3), with $\varphi_0(x, y) = 20 \cdot 2\pi \cdot \sqrt{x^2 + y^2}$ and $\{(x_i, y_i)\} = \{(0.2, -0.25), (-0.2, -0.37), (0.0, 0.2), (-0.25, 0.3), (0.2, 0.43)\}$.



Fingerprint Sample Synthesis. Figure 4 Sherlock and Monro model: each element of the orientation field is considered as a complex number.



Fingerprint Sample Synthesis. Figure 5 An example of Arch (**a**), Tented Arch (**b**), Right Loop (**c**) and Whorl (**d**) orientation field as generated by the Sherlock and Monro model. In (**e**), an example of left-loop orientation field superimposed to a real left-loop fingerprint with coincident singularity positions.

introduces more degrees of freedom to cope with the orientation variability that may characterize orientation fields with coincident singularities. The orientation θ at each point $\mathbf{z}$ is calculated as:

$$\theta = \frac{1}{2}\left[\sum_{i=1}^{n_d} g_{\mathbf{ds}_i}(\arg(\mathbf{z}-\mathbf{ds}_i)) - \sum_{i=1}^{n_c} g_{\mathbf{ls}_i}(\arg(\mathbf{z}-\mathbf{ls}_i))\right] \tag{5}$$

where $g_k(\alpha)$, for $k \in \{ls_1, \ldots, ls_{n_c}, ds_1, \ldots, ds_{n_d}\}$, are piecewise linear functions capable of locally correcting the orientation field with respect to the value given by Sherlock and Monroe model:

$$g_k(\alpha) = \bar{g}_k(\alpha_i) + \frac{\alpha - \alpha_i}{2\pi/L}(\bar{g}_k(\alpha_{i+1}) - \bar{g}_k(\alpha_i)) \tag{6}$$

for $\alpha_i \le \alpha \le \alpha_{i+1}$, $\alpha_i = -\pi + \frac{2\pi i}{L}$.

Each function $g_k(\alpha)$ is defined by the set of values $\{\bar{g}_k(\alpha_i) | i = 0..L-1\}$, where each value is the amount of correction of the orientation field at a given angle (in a set of L angles uniformly distributed between $-\pi$ and π). If $\bar{g}_k(\alpha_i) = \alpha_i \forall i \in \{0..L-1\}$ (i.e. $g_k(\alpha)$ is the identity function), the model coincides with that of Sherlock and Monro.

Figure 6a and b show two examples of orientation fields generated according to the Vizcaya and Gerhardt model; these images are definitely more realistic than those in Fig. 5. The superiority of the Vizcaya and Gerhardt model in approximating existing ridge patterns is also evident from the comparison between Fig. 6c and d.

In 2000, Cappelli et al. introduced a ridge pattern generation approach based on the following steps [12]:

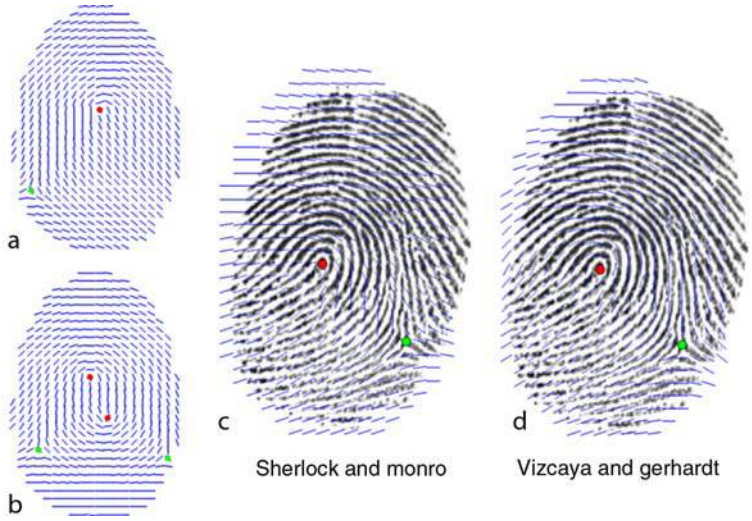
1. Orientation field generation
2. Frequency map generation
3. Ridge pattern generation

Step 1 adopts the Vizcaya and Gerhardt model for generating the orientation field starting from the positions of loops and deltas; for generating arch type patterns (which do not contain any singularity), a simple sinusoidal function, whose frequency and amplitude are tuned to control the arch curvature and aspect, is used.

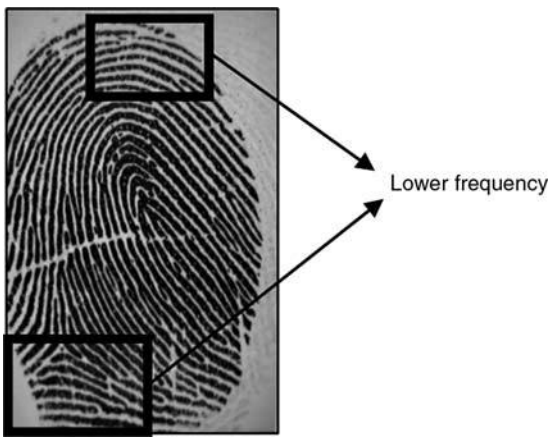
Step 2 creates a frequency map (see Fingerprint Feature Extraction) on the basis of some heuristic criteria inferred by the visual inspection of a large number of real fingerprints (for instance, in the regions above the northernmost loop and below the southernmost delta, the ridge-line frequency is often lower than in the rest of the fingerprint, see Fig. 7).

Finally step 3, given an orientation field and a frequency map as input, generates a ridge line pattern by iteratively enhancing an initial image (containing one or more isolated points) through ▶ Gabor filters. The filters are applied at each pixel (x, y) and adjusted according to the local ridge orientation ϕ_{xy} and frequency v_{xy}:

$$gabor\left(r, s : \phi_{xy}, v_{xy}\right) = e^{-\frac{(r+s)^2}{2\sigma^2}} \cdot \cos\left[2\pi v_{xy}\left(r\sin\phi_{xy} + s\cos\phi_{xy}\right)\right] \tag{7}$$



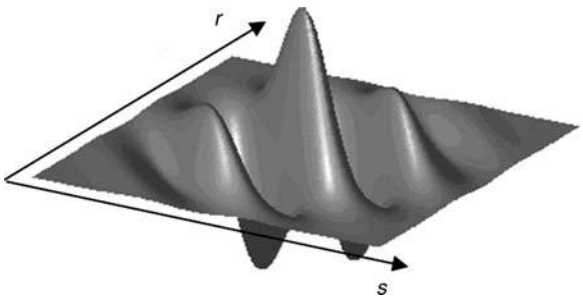
Fingerprint Sample Synthesis. **Figure 6** An example of Right Loop (**a**) and Whorl (**b**) orientation fields, as generated by the Vizcaya and Gerhardt model. In (**c**) and (**d**) the orientation fields produced by the two models, for a given fingerprint, are compared.



Fingerprint Sample Synthesis. **Figure 7** An example of a right-loop fingerprint where the ridge-line frequency is lower in the regions above the loop and below the delta.

Parameter σ, which determines the bandwidth of the filter, is set according to the frequency, so that the filter does not contain more than three effective peaks (see Fig. 8).

While one could reasonably expect that iteratively applying "striped" filters to random images would simply produce striped images, very realistic minutiae are generated at random positions. Based on their experiments, in [12] the authors argue that minutiae primarily originate from the ridge-line disparity produced by local convergence/divergence of the

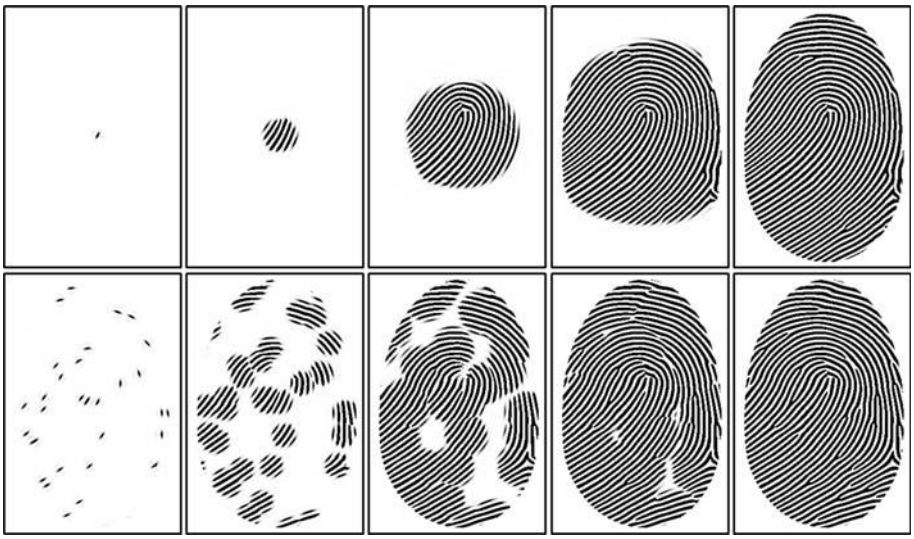


Fingerprint Sample Synthesis. **Figure 8** An example of Gabor filter used in step 3: note that the bandwidth is adjusted so that the filter does not contain more than three peaks.

orientation field and by frequency changes. In Fig. 9, examples of the iterative ridge-line generation process are shown; the authors experimentally found that increasing the number of initial points determines a more irregular ridge pattern richer of minutiae: this is not surprising, since expanding distinct image regions causes interference where regions merge, thus favoring the creation of minutiae (see Fig. 10).

Generation of Synthetic Fingerprint Impressions

Several factors contribute in making the impressions of a real finger substantially different when captured



Fingerprint Sample Synthesis. **Figure 9** Some intermediate steps of a fingerprint-generation process starting from a single central point (top) and from a number of randomly located points (bottom). Usually, increasing the number of initial points determines a more irregular ridge pattern richer of minutiae.

Fingerprint Sample Synthesis. **Figure 10** Genesis of a minutia point during the merging of the two regions originated by two different initial points.

by an on-line acquisition sensor (see ▶ Fingerprint Device):

1. Displacement in x and y direction and rotation
2. Different touching areas
3. Non-linear distortions produced by non-orthogonal pressure of the finger against the sensor
4. Variations in the ridge-line thickness given by pressure intensity or by skin dampness
5. Small cuts or abrasions on the fingertip
6. Background noise and other random noise

In 2002, Cappelli et al. proposed an evolution of the approach introduced in [13], which is able to simulate most of the above factors, thus generating very realistic fingerprint impressions. Starting from a synthetic ridge-line pattern, the main steps involved in the simulation of a fingerprint impression are: (1) Variation of the ridge thickness; (2) Skin distortion; (3) Noising and global translation/rotation; (4) Background generation. The subsections that follow briefly describe the various steps, as they were proposed by Cappelli [14].

Variation of the Ridge Thickness

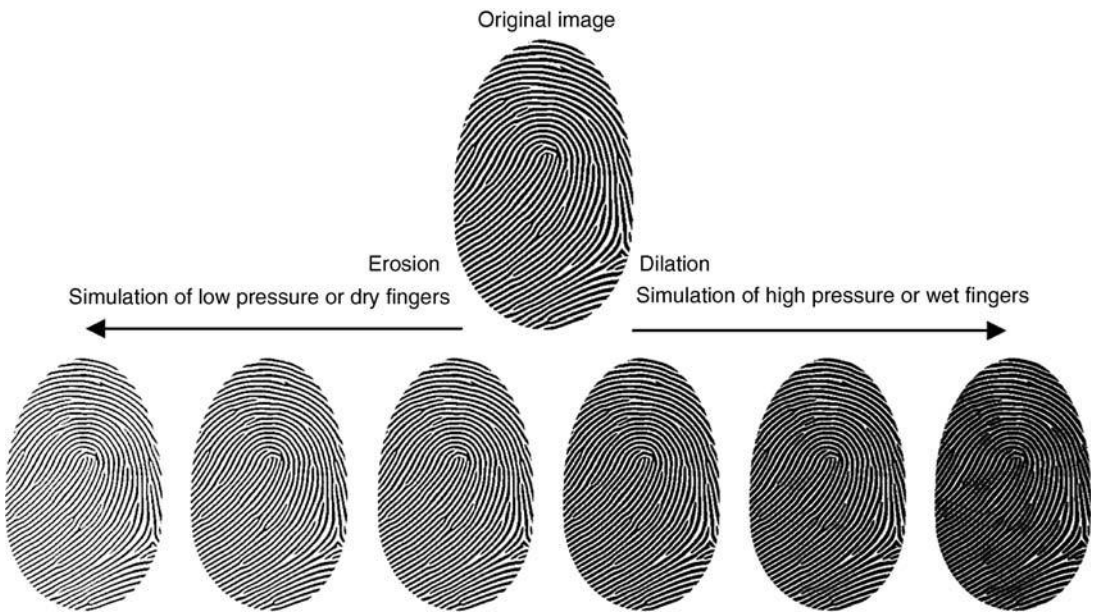
Skin dampness and finger pressure against the sensor platen have similar effects on the acquired images: when the skin is dry or the pressure is low, ridges appear thinner, whereas, when the skin is wet or the pressure is high, ridges appear thicker (see Fig. 11).

Morphological operators (see Image Preprocessing) are applied to the ridge line pattern, to simulate different degrees of dampness/pressure. In particular, the erosion operator is applied to simulate low pressure or dry skin, while the dilation operator is adopted to simulate high pressure or wet skin (see Fig. 12).

Skin Distortion

One of the main aspects that distinguish the different impressions of the same finger is the presence of non-linear distortions, mainly due to skin deformations

Fingerprint Sample Synthesis. **Figure 11** Three impressions of the same real finger as captured when the finger is dry, normal and wet, respectively.

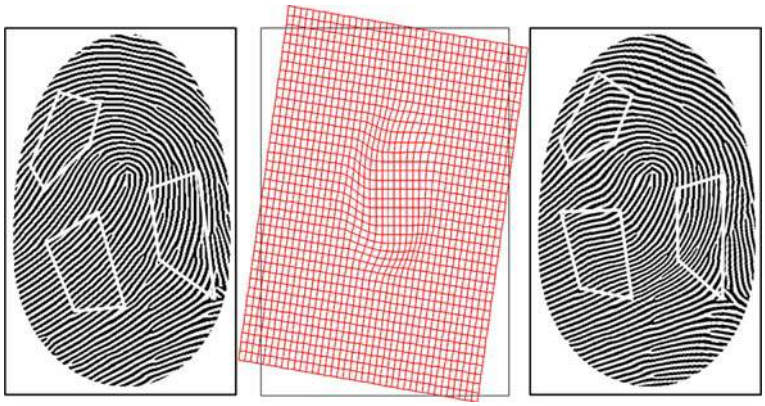


Fingerprint Sample Synthesis. **Figure 12** The application of different levels of erosion/dilation to the same ridge line pattern.

according to different finger placements over the sensing element (see Fig. 13). In fact, due to the skin plasticity, the application of forces, some of whose components are not orthogonal to the sensor surface, produces non-linear distortions (compression or stretching) in the acquired fingerprints (see ▶ Fingerprint Matching).

In "Synthetic Fingerprint Generation" [14], the skin-distortion model introduced by Cappelli, Maio, and Maltoni [15] is exploited. While in the latter, the distortion model was applied to re-map minutiae points, in order to improve fingerprint matching,

Fingerprint Sample Synthesis. **Figure 13** Two impressions of the same real finger where a few corresponding minutiae are marked to highlight distortion.



Fingerprint Sample Synthesis. Figure 14 A synthetic ridge line pattern (on the left) and a distorted impression (on the right); the equivalent distortion of a square mesh is shown in the middle. To better highlight the non-linear deformations, some corresponding minutiae are connected by white segments in both the fingerprint images.

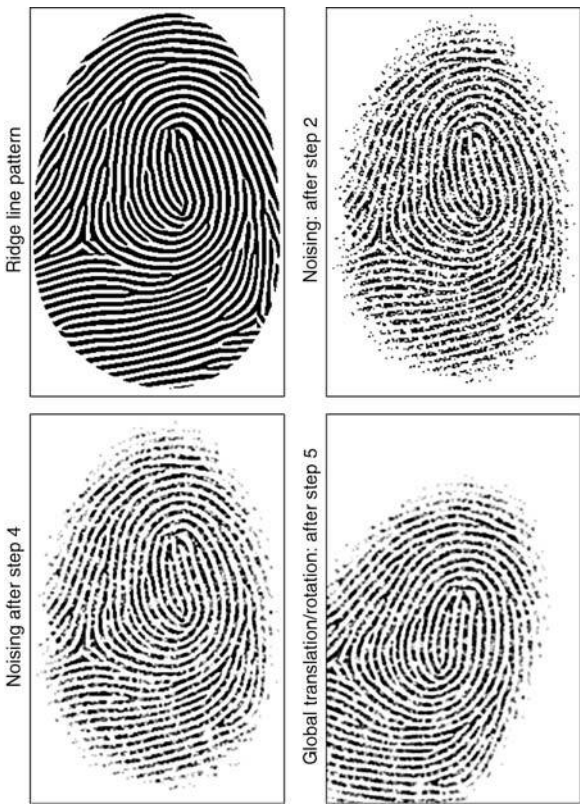
here the mapping has to be applied to the whole image, in order to simulate realistic distorted impressions. In Fig. 14, a ridge line pattern and its distorted impression are shown.

Noising and Global Translation/Rotation

During fingerprint acquisition, several issues contribute to deteriorate the original signal, thus producing a gray-scale noisy image: irregularity of the ridges and their different contact with the sensor surface, presence of small pores within the ridges, presence of very-small-prominence ridges, gaps and clattering noise due to non-uniform pressure of the finger against the sensor. Furthermore, the fingerprint is usually not perfectly centered in the image and can present a certain amount of rotation. The noising phase sequentially performs the following steps:

1. Isolate the valley white pixels into a separate layer. This is simply performed by copying the pixels brighter than a fixed threshold to a temporary image
2. Add noise in the form of small white blobs of variable size and shape. The amount of noise increases with the inverse of the fingerprint border distance
3. Smooth the resulting image with a 3×3 averaging box filter
4. Superimpose the valley layer to the image obtained
5. Rotate and translate the image

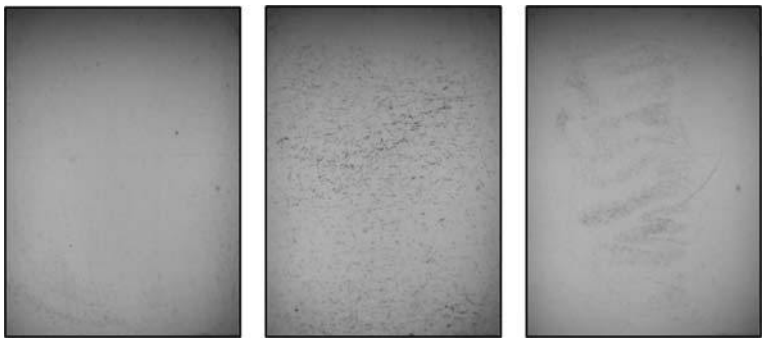
Steps 1 and 4 are necessary to avoid an excessive overall image smoothing. Figure 15 shows an example where the intermediate images produced after steps 2, 4 and 5 are reported.



Fingerprint Sample Synthesis. Figure 15 An example of noising and global translation/rotation, where the intermediate images produced after steps 2, 4 and 5 are reported.

Background Generation

The output of the previous step is a fingerprint that appears realistic, but the image background is completely white. In order to generate backgrounds



Fingerprint Sample Synthesis. **Figure 16** Examples of background-only images (acquired from an optical sensor) used for training the background generator.

Fingerprint Sample Synthesis. **Figure 17** Three synthetic images with backgrounds generated according to the model in [14].

similar to those of fingerprints images acquired with a given sensor, a statistical model based on the KL transform (see ► Dimensionality Reduction) is adopted. The model requires a set of background-only images as a training set (see Fig. 16): a linear subspace that represents the main variations in the training backgrounds is calculated and then used to randomly generate new backgrounds.

Figure 16 shows some examples of the background images (obtained from an optical acquisition sensor) used as a training set for the background generation step; Fig. 17 reports three synthetic fingerprints with backgrounds generated according to the above-described model.

Related Entries

► Anatomy of Fingerprint
► Biometric Sample Synthesis
► Fingerprint Classification
► Fingerprint Databases and Evaluation
► Fingerprint Features
► Fingerprint Singularities, Minutiae, Pores

References

1. UK Biometrics Working Group: Biometric Evaluation Methodology (2002)
2. Biometric Information Management and Security, American National Standards Institute, X9.84, 2001
3. Cappelli, R., Maio, D., Maltoni, D., Wayman, J.L., Jain, A.K.: Performance evaluation of fingerprint verification systems. IEEE Trans. Pattern Anal. Mach. Intell. **28**(1), 3–18 (2006)
4. Kücken, M.: Models for fingerprint pattern formation. Forensic Sci. Int. **171**(2–3), 85–96 (2007)
5. Sherstinsky, A., Picard, R.W.: Restoration and enhancement of fingerprint images using M-lattice – a novel non-linear dynamical system. Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem, pp. 195–200 (1994)
6. Kücken, M., Newell, A.C.: A model for fingerprint formation. Europhys. Lett. **68**(1), 141 (2004)

7. Cummins, H.: Epidermal-ridge configurations in developmental defects, with particular reference to the ontogenetic factors which condition ridge direction. Am. J. Anat. **38**, 89–151 (1926)
8. Bonnevie, K.: Studies on papillary patterns in human fingers. J. Genet. **15**, 1–111 (1924)
9. Kosz, D.: New numerical methods of fingerprint recognition based on mathematical description of arrangement of dermatoglyphics and creation of minutiae. In: Mintie, D. (ed.) Biometrics in Human Service User Group Newsletter, (1999)
10. Bicz, W.: "The idea of description (reconstruction) of fingerprints with mathematical algorithms and history of the development of this idea at Optel," (Optel, 2003). http://www.optel.pl/article/english/idea.htm. Accessed 18 Dec 2007
11. Sherlock, B.G., Monro, D.M.: A model for interpreting fingerprint topology. Pattern Recognit. **26**(7), 1047–1055 (1993)
12. Cappelli, R., Erol, A., Maio, D., Maltoni, D.: Synthetic fingerprint-image generation. Proceedings of the 15th International Conference on Pattern Recognition, vol. 3, pp. 475–478. Barcelona (2000)
13. Cappelli, R., Maio, D., Maltoni, D.: "Synthetic fingerprint-database generation." Proceedings of the 16th International Conference on Pattern Recognition, vol.3, pp. 744–747. Québec City (2002)
14. Cappelli, R.: Synthetic fingerprint generation. In: Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S. (eds.) Handbook of Fingerprint Recognition. Springer, New York (2003)
15. Cappelli, R., Maio, D., and Maltoni, D.: "Modelling plastic distortion in fingerprint images". Proceedings of the Second International Conference on Advances in Pattern Recognition (ICAPR2001), Rio de Janeiro, pp. 369–376 (2001)

Fingerprint Scan

► Biometric Sample Acquisition

Fingerprint Sensor

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Fingerprint Signatures

► Fingerprint Features

Fingerprint Singularity

Fingerprint singularity is defined as a core or delta of a fingerprint. Minutiae have dual correspondence between the normal image and density-inverted image, that is, terminations appear as bifurcations and vice versa. However, singularity does not have such trait.

► Fingerprint Image Enhancement

Fingerprint Skeletonization

Fingerprint skeletonization, also referred to as thinning, is the process of reducing the width of binarized ridgelines to 1 pixel. Standard thinning algorithms are applicable. Modified methods based on local ridge orientation have been proposed to improve skeletonization accuracy. Post-processing for skeleton image, such as skeleton adjustment, is also important.

► Fingerprint Image Enhancement

Fingerprint Templates

Wei-Yun Yau
Institute for Infocomm Research, Agency for Science, Technology & Research, Singapore

Synonym

Fingerprint representation

Definition

A fingerprint template is a set of stored fingerprint features extracted from the fingerprint of a user. It is stored during the enrollment process to represent the actual owner of the fingerprint. It is subsequently compared directly to the fingerprint features of the query fingerprint in order to establish whether the

query fingerprint is obtained from the same person as the actual owner. It should be noted that the original fingerprint or its enhanced or compressed form is not a fingerprint template.

Introduction

As discussed in the section on general biometrics, the operation of a fingerprint recognition system, just like any other biometric system, follows a common process flow as shown in Fig. 1.

A fingerprint sensor is required to capture the fingerprint image which is then processed by a feature extractor to obtain the unique features of the fingerprint. If the user is new, the features extracted are stored in a database, typically along with other personal details of the new user such as name, and identification number. This set which stores the fingerprint feature is commonly known as a *fingerprint template*. The process by which this is done is called the *enrollment* process. Subsequently, when the user wants to use the fingerprint recognition system, the fingerprint features extracted from the fingerprint image acquired live or as a query image provided into the system are compared against the stored *fingerprint template(s)*. If the comparison involves only one *fingerprint template* from the database, such as when the user key in the name to retrieve the enrolled fingerprint template, the comparison process is called *verification*. Alternatively, the comparison can be done against all the fingerprint templates stored in the database and such a process is referred to as *identification*.

Composition of Fingerprint Template

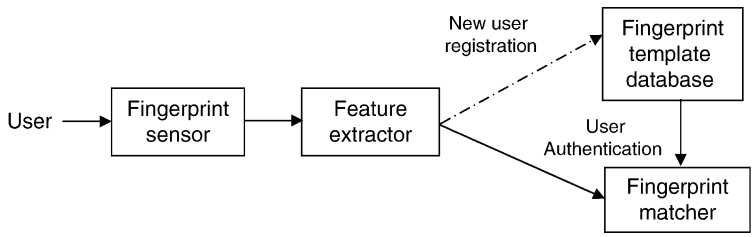
In general, a fingerprint template contains the unique features extracted from the fingerprint image. However, the exact content varies according to the type of algorithm used to extract and match the fingerprint. Nevertheless, if the stored file is merely the enhanced or compressed fingerprint image or the original fingerprint image itself, it is not considered a fingerprint template. There are two general types of algorithm used in fingerprint feature extraction and matching, namely, minutia-based and pattern-based or ridge feature-based [1, 2]. Minutia arises when a fingerprint ridge comes to an end (called ridge ending) or when it forks out into two ridges (called bifurcation). A sample fingerprint image with the detected minutiae is presented in Fig. 2. Minutia detection is a complex process and is thus beyond the purview of this contribution.

Each minutia, F, can be represented by a parameter vector $F = (x, y, \varphi, t)^T$ where (x, y) is the coordinate in the image, φ the local ridge direction and t the type of the minutia (i.e., bifurcation or ridge ending). The basic composition of the minutia template, $\mathbf{S}$, of a fingerprint image is then the set of all n valid minutia parameter vectors found in the fingerprint image given by:

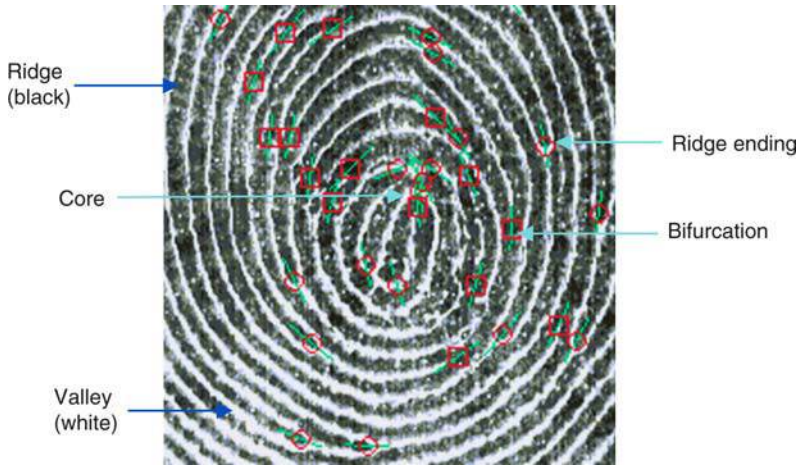
$$\mathbf{S} = F_k = \left(x_k, y_k, \varphi_k, t_k\right), k = 1, 2, \ldots, n \quad (1)$$

Apart from the minutia information, the ridge count [2] between two minutiae, which is the number of ridges intersecting a straight line joining two minutiae, is commonly used and included in the template. The non-minutia data commonly extracted and included in the template are the location, direction and the number of core and ▶ delta points. There are many other details that can be extracted and included in the template such as a short ridge line information associated to the minutia, and the number and type of minutia encountered by the straight line used in the ridge count with the aim to improve the performance of fingerprint matching.

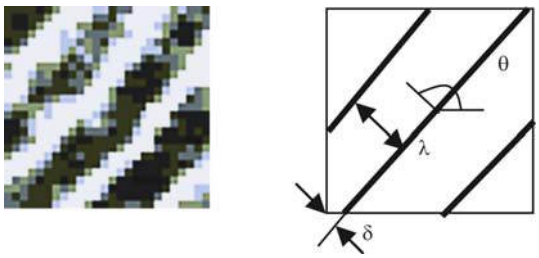
A popular pattern-based approach is the FingerCode [3] approach. The fingerprint image is tessellated



Fingerprint Templates. Figure 1 Process Flow of a Fingerprint Recognition System.



Fingerprint Templates. **Figure 2** A Fingerprint Image Showing Detected Ridge Endings and Bifurcations.



Fingerprint Templates. **Figure 3** A Sample Set of Spectral Triplets Representation (right) of a Fingerprint Image Block (left).

into sectors and bands with respect to a reference point, such as the ► core point. A bank of Gabor filters is then applied to each cell in the tessellation, bounded by the boundary delineated by the sector and band of the tessellation. For every cell, the average absolute deviation of each filter response over all the pixels in the cell is computed and used as an element of the feature vector, called the FingerCode. To approximate rotation invariant, rotate the FingerCode cyclically. Thus, the combined FingerCode becomes the template for this approach. Another proposed approach includes dividing the fingerprint image into small blocks. For each block, the spectral information that describes the fingerprint pattern in that block as closely as possible is obtained. This can be done using Discrete Fourier Transform, Gabor Filterbanks or by selecting the spectral component from a predefined set of spectral triplets [4] (see Fig. 3). The parameters describing the spectral component for each block, such as $(\theta, \lambda, \delta)$ for the spectral triplets, is quantized to limited discrete values and then stored as a feature vector. The set of feature vectors for all the blocks in the region of interest of the fingerprint image is then stored as the fingerprint template.

Storage of Fingerprint Template

Usually fingerprint templates are stored in a central database residing in a central database server. To perform a match, the extracted features from the query fingerprint image are sent to the server. Such a model requires connection from the point where the query image is acquired to the central server. Since the fingerprint template is stored in a central server, it carries a notion of "big brother" which is a cause of concern for the privacy advocates as such a system is capable of tracking an individual. Another model for the storage of fingerprint template is the distributed database concept. The fingerprint templates are stored at each unit where the use of the fingerprint system is the most common. However, if a user wishes to be recognized in the other system, the fingerprint template has to be sent to the unit in advance, usually via a central server which acts as a backup and synchronization unit. Such a model is usually preferred to the central database model if the usage spans over a large geographical area. Instead of depending on a database, the fingerprint template can also be stored in a token such as a smartcard, memory stick or thumb drive in a fully decentralized model. A smartcard is usually preferred since it is generally regarded as more secure. The user carries the token containing his or her fingerprint

template with him/her. To perform fingerprint matching, the user has to present the token and the fingerprint template stored in the token is retrieved for matching with the query features. Alternatively, the matching is performed inside the token itself with the query features sent into the token. As such, the use of the fingerprint system is not dependent on any connectivity. Also, the number of users can easily be scaled when managing the template database. Unfortunately, once the token is lost, the genuine user is unable to use the system.

Template Synthesis

Various types of fingerprint sensors are available that capture live fingerprint images. Traditionally, those used for law enforcement purposes require a sensor with a sensing area of at least 2.54cm × 2.54cm. The consumer version is usually smaller, about a quarter of the size or even smaller. However, such a small sensor is not able to image the complete portion of the finger that touches the sensor. Consequently, if the user does not position his/her finger such that the contact portion of the skin is largely similar to the portion used during enrollment, then the matching will fail. This will result in false non-match, causing inconvenience to the user. To solve this, an image **mosaic**king technique has been proposed [5] for constructing a composite fingerprint from an image sequence of partial fingerprints. This is done by applying a low pass filter to smooth the images and then compressing the intensity to the range of [10, 20]. The images are then aligned using the Iterative Closest Point algorithm [6] before superimposing the aligned images to form the ▶ mosaic. Another **mosaic**king technique for rolled fingerprints has been proposed in [7]. Alternatively, a minutia-based template synthesis approach to combine the various fingerprint templates obtained from the small fingerprint images into a composite template which resembles the template obtained using a larger fingerprint image has been developed in [8].

Minutia-based template synthesis is performed by finding all the correspondent or matched minutiae between two fingerprint images, I_R and I_1, to be synthesized. Based on the matched minutiae, an affine transformation that maps the minutiae from I_1 to I_R is then determined. This is repeated until all the other fingerprint images are synthesized. The experimental comparison [9] revealed that the template synthesis approach is faster, less affected by elastic deformation, and is more suitable for larger partial images while the image **mosaic**king approach is more appropriate when accurate performance for small partial images is required.

Template Improvement

Fingerprint images are often corrupted by noise, imaging artifacts and affected by the skin condition (wet, dry), amount of pressure exerted when touching the sensor, etc. This causes the occurrence of missing minutiae (valid minutiae are not detected) or spurious minutiae (false minutiae detected). Thus, accurate detection of minutia is a very challenging task. If many dropped or spurious minutiae are present in the fingerprint template, the usability of the fingerprint recognition system is affected. To expect the user to re-enroll regularly may cause a lot of convenience. The purpose of template improvement is to improve the fingerprint template using multiple fingerprint images captured over a period of time [10]. For each minutia in the template, it is initialized with a default certainty level. When a query fingerprint submitted after a time interval is matched above a predefined threshold, the template and the certainty level associated with each minutia will be updated. This is done by finding those unmatched minutiae in the template within the region which overlaps with the query fingerprint and then reducing their certainty level by a predefined weight, α. Next, all the unmatched minutiae found in the query fingerprint outside of the overlapping region will be included in the template using the template synthesis technique but with a reduced certainty level of $(1 - \alpha)$. Then all minutiae in the template with a certainty level lower than a predefined threshold will be removed. In this way, spurious minutiae can be eliminated after many unmatched iterations while the missing minutiae can be incorporated.

Template Interchange

There are many ways in which a fingerprint system can generate a fingerprint template. In order to facilitate the interchanging of fingerprint templates among the

various fingerprint systems from different vendors, fingerprint templates have to be coded in a consistent manner. This is defined by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). The ISO/IEC documents specify the standards for the minutia-based template [11], the pattern-based template [4], and the combined minutia and pattern template, called the pattern skeletal [12]. The standards specifying consistent formats for the construction of fingerprint templates comprise 3 sections:

1. A header which describes the generic information about the template and its source.
2. A normative section which describes all the mandatory features to be included and the manner in which they have to be coded.
3. A non-normative section which allows for other non-mandatory information to be included in the fingerprint template.

Image Reconstruction from Template

It is often assumed that the minutiae-based template cannot be used to construct back the corresponding original fingerprint, partly because the way the data is stored in the template is proprietary. However, it has been shown that this is possible with a standard template [13], such as those defined by the ISO/IEC [9]. The general idea is to reconstruct the orientation pattern using the orientation modeling approach [14] and the fingerprint area based on the template information. Subsequently, the ridge pattern is developed by applying a high gain Gabor filter adjusted to the local frequency and orientation and then rendering it to make the image look realistic [2].

Summary

A fingerprint template contains the unique features of a fingerprint image and can be used for fingerprint matching. The exact composition of the template is dependent on the algorithm used to extract the unique features. Nevertheless, international standards exist to facilitate the interchanging of the template. It can be stored in a database which can either be centrally managed, distributed or stored in portable tokens instead of a database. Template synthesis and template improvement techniques can be used to improve the performance of the system when dealing with small fingerprint sensor and poorly enrolled fingerprint templates respectively.

Related Entries

- ▶ Enrolment
- ▶ Fingerprint matching
- ▶ Minutia

References

1. Jain, A.K., Hong, L., Pankanti, S., Bolle, R.: An Identity-authentication System Using Fingerprints. Proceedings of the IEEE **85**(9), 1365–1388 (1997)
2. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer-Verlag, New York (2003)
3. Jain, K., Prabhakar, S., Hong, L., Pankanti, S.: Filterbank-based Fingerprint Matching. IEEE Trans. Image Process. **9**(5), 846–859 (2000)
4. Standards document ISO/IEC 19794–3: Biometric Data Interchange Formats – Part 3: Finger Pattern Spectral Data
5. Jain, A.K., Ross, A.: Fingerprint Mosaicking, In: Proceedings of the International Conference on Acoustic, Speech Signal Processing, vol. 4, pp. 4064–4067. Washington, DC (2002)
6. Besl, P.J., McKay, N.D.: A Method for Registration of 3D Shapes, IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
7. Bolle, R.M., Ratha, N.K., Connell, J.H.: Image Mosaicking for Rolled Fingerprint Construction, In: Proceedings of International Conference on Pattern Recognition, vol. 2, pp. 1651–1653 (1998)
8. Yau, W.Y., Toh, K.A., Jiang, X.D., Chen, T.P., Lu, J.W.: On Fingerprint Template Synthesis, In: Proceedings of the 6th International Conference on Control, Automation, Robotics and Vision 5–8 (2000)
9. Moon, Y.S., Yeung, H.W., Chan, K.C., Chan, S.O.: Template Synthesis and Image Mosaicking for Fingerprint Registration: An Experimental Study, In: Proceedings of the IEEE Conference on Acoustics, Speech, Signal, vol. 5, pp. V-409–V-412 (2004)
10. Jiang, X., Ser, W.: Online Fingerprint Template Improvement, IEEE Trans. Pattern Anal. Mach. Intell. **24**(8), 1121–1126 (2002)
11. Standards document ISO/IEC 19794–2: Biometric Data Interchange Formats – Part 2: Finger Minutiae Data
12. Standards document ISO/IEC 19794–8: Biometric Data Interchange Formats – Part 8: Finger Pattern Skeletal Data

13. Cappelli, R., Lumini, A., Maio, D., Maltoni, D.: Fingerprint Image Reconstruction from Standard Templates, IEEE Trans. Pattern Anal. Mach. Intell. **29**(9), 1489–1503 (2007)
14. Vizcaya, P., Gerhardt, L.: A Nonlinear Orientation Model for Global Description of Fingerprints, Pattern Recognit. **29**(7), 1221–1231 (1996)

Fingerprint Thinning

Obtaining a 1 pixel wide digital skeleton of ridges.

► Fingerprint Features

Fingerprint, Forensic Evidence of

Didier Meuwly
Netherlands Forensic Institute, The Hague, The Netherlands

Synonyms

Fingermark identification procedure; Automatic fingerprint identification system; Forensic evaluation of fingerprints and fingermarks.

Definition

Forensic evidence of fingerprint is the field of forensic expertise related to the inference of the identity of source from the examination of all the friction ridge skin, namely the fingers, the palms, the toes, the soles, and their marks. But for the sake of simplicity, the text is mainly focused on fingerprints and fingermarks. The extreme variability of the fingerprints derives firstly from the knowledge of the morphogenesis of the papillary ridges pertaining to embryology and, secondly, from statistical researches pertaining to dactyloscopy. This variability is mainly used in four different processes within forensic science: identity verification, forensic intelligence, forensic investigation, and forensic evaluation. The first three processes are based on the use of Automatic Fingerprint Identification Systems (AFIS). The fourth process, forensic evaluation, is an expert-based process, built on procedure, training, and experience. The procedure and practice vary a lot between countries, principally regarding the threshold used for forensic identification. Most of the European and South American countries favor a quantitative approach based on a numerical standard when the USA, UK, and most of the Scandinavian countries have adopted a qualitative approach based on the experience and knowledge of the dactyloscopist. For both approaches, the decision is an expert opinion that is deterministic: exclusion, inconclusive or identification. As the current practice is not error-free and partly based on the subjective probabilities of the dactyloscopists, efforts are made to develop a new approach based on a logical inference model and statistical probabilities, in order to assist the dactyloscopists in producing a logical, testable, and quantitative evaluation of the fingerprint evidence.

Nomenclature

At the end of the 19th century William Herschel and Henry Faulds expressed the principles of the forensic use of fingerprints and fingermarks: the use of fingerprints and fingerprint databases for the identification of serial offenders and the use of fingermarks to establish a link between a crime scene or an object and an individual. In literature, confusion exists between the term fingerprint and fingermark. This article uses a uniform terminology: the finger dermatoglyphics and their standard rolled inked impressions are named fingerprints, whereas recovered traces left by unprotected fingers are named fingermarks. In criminal records, reference prints are collected using forms named ten-print cards.

Individuality of the Fingerprint

Confusion surrounds the terms *identity*, *identify*, and *identification* in forensic science. This is clearly demonstrated in popular practice, when the perpetrator of an infringement is said to be "identified from her/his fingerprints". The perpetrator is not identified, but individualized. What is proved by the fingerprints is individuality. To individualize a human being on the basis of fingermarks in forensic science ultimately consists in determining if an individual is the source of

the fingermark linked to the criminal activity [1]. The individuality of fingerprints derives firstly from the knowledge of the morphogenesis of the papillary ridges pertaining to embryology and, secondly, from statistical researches pertaining to dactyloscopy.

Morphogenesis

The friction ridge skin morphogenesis offers a biological basis to explain the variability in friction ridge patterns. The morphogenesis of the human hands and feet starts during the 6th week of the estimated gestational age (EGA). The pattern of ridge skin is established from the 10th week to the 14th week of EGA when the basal layer of the volar epidermis becomes folded and forms the primary ridges. This process is influenced by the volar pads, local eminences of subcutaneous tissue in well-defined locations of the volar surfaces. It is conjectured that the inversion of the volar pads creates tensions in the epidermis that align the ridge pattern [2]. From this moment on up to the 16th week of EGA, the tissues growing under the dermis, named volar pads, induce physical stress in the cell layers constituting this dermis. This physical stress forms a two-dimensional structure of ridges on the palms, the soles, the fingers tips, and the toes. From the 16th to the 24th week of EGA, the dermis matures; secondary dermal ridges start to develop between the primary dermal ridges and bridges, named dermis papillae, appear between the apex of the primary and secondary ridges. After 24 weeks of EGA, the development of the dermis is finalized and the epidermis is gradually formed by cell development from the dermis, named papillary ridges. In its final stage, the papillary ridges grow as a three-dimensional structure based on the two-dimensional pattern. The anchorage of this epidermal structure in the dermis ensures the stability and the permanence of the dermatoglyphics. Therefore a permanent modification or destruction of the dermatoglyphics can only occur in case of destruction of the dermis [3].

Variability of the Fingerprint

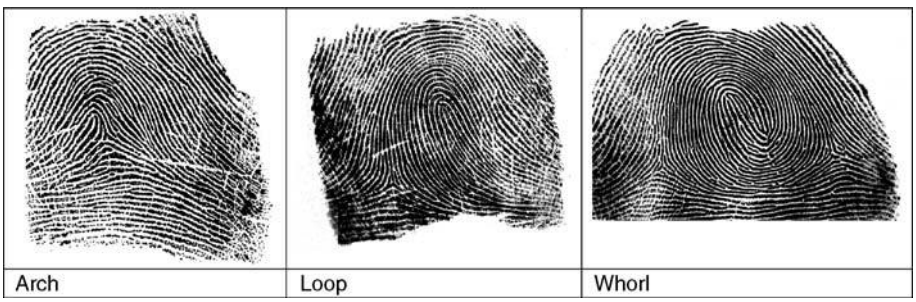
The fingerprint is expressed through the interaction of genotype, development, and environment; therefore this biometric modality is qualified as epigenetic, similar to the iris of the eye but contrarily to a DNA sequence, from which by instance a DNA profile is extracted, that is genetically determined. The information content in the fingerprint ridges is structured in three levels named the general pattern, the minutiae, and the third level details.

General Pattern

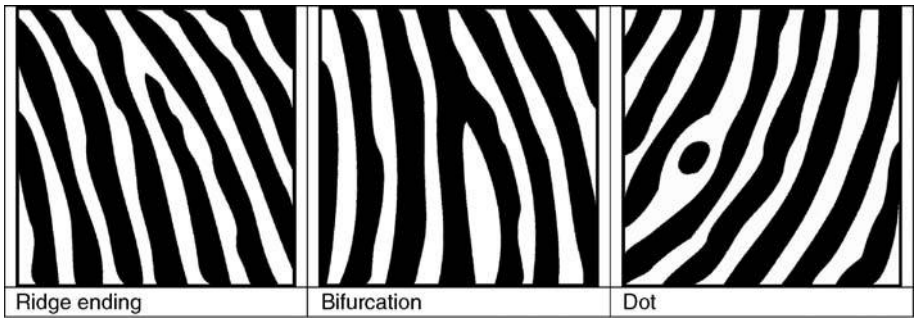
The general shape of the ridge flow, named general pattern, is to some degree indirectly genetically inherited and is classified in three generic types: arches (simple or tented), loops (left or right), and whorls (including various composite forms). The approximate center of the general pattern is named the core, and the small area where 3 flows of ridges meet to form a triangular pattern is called a delta. Arches have no delta, loops have 1 delta, and whorls, 2 deltas (Fig. 1).

Minutiae

In addition to the general ridge flow of the ridges, deviations appear along the papillary ridges. They are named minutiae, and can be classified in three basic types: ridge ending, bifurcation, and dot (Fig. 2). All other denominations, employed at the convenience of



Fingerprint, Forensic Evidence of. **Figure 1** Examples of fingerprint with different general patterns: arch, loop, and whorl.



Fingerprint, Forensic Evidence of. **Figure 2** The three basic structures of minutiae: the ridge ending, the bifurcation, and the dot.

the users or for statistical purposes, are a combination of two or three minutiae of basic type.

The minutiae contribute the most to the selectivity of the fingerprint, due to the combination of their spatial arrangement along the ridges and their intrinsic characteristics: type, location, and orientation. The selectivity offered by a minutiae configuration present on a fingerprint or on a fingermark is a function of their number, type, and topology (relative position and orientation on the ridges).

The process underlying the development of the minutiae is not known yet, but models offered by mathematical biology and empirical studies suggest that it is epigenetic [2]. For ridge endings, bifurcations, and dots, more correlations are observed on fingerprints of monozygotic twins as opposed to dizygotic twins [4]. Correlations are also observed between the number of minutiae and the finger number, which can be explained by the fact that the surface of the fingertip of the thumb is bigger than the surface of the fingertip of the little finger. The relative frequencies of the minutiae type are correlated with gender, but no difference has been observed between the fingerprint characteristics of the left and right hands [5].

Third Level Details

The study of the friction ridge details may be further subdivided into the description of ridge contours or edges, and the position and the shape of the pores [6]. However, the degree of agreement between dactyloscopists on the value of these latter characteristics is limited so far, and no systematic study supports the different opinions.

Statistical Research

The first statistical investigations were conducted at the end of the nineteenth and at the beginning of the twentieth century, but the initial models were developed on the basis of unrealistic premises: it was presumed that each minutiae type appeared with the same probability and independently of each other on the ridge skin surface. More sophisticated models were developed later during the twentieth century, first including the unbalance between the minutiae type (e.g., the bifurcations are more rare than the ridge endings) and then including the uneven density of the minutiae (e.g., the density of minutiae increases in the centre and delta zones) [7].

Statistical studies mainly focus on the second level features and especially the spatial arrangement of minutiae, while studies of other fingerprint features remain too seldom. These studies on minutiae provide extremely valuable fundamental knowledge about the degree of randomness of minutiae configurations, but they cannot be used yet for the deployment of large-scale, case-specific statistical evaluation of the fingermark evidence. Current statistical models simplify reality, emphasizing the statistical behavior of minutiae, and adopting a restricted view of the overall factors like the general pattern, the main ridge flows, the ridge edges, or the pores. Nevertheless, this new approach aims to offer a uniform framework and a transparent methodology to the dactyloscopists. Coupled to a logical inference model originating in the Bayes theorem, these models aim to assist them in producing a logical, testable, and quantitative evaluation of the fingerprint evidence based on statistical probabilities [8].

Classification of Fingerprints and Fingermarks

Manual Classification

For about a century the classification of fingerprints based on general patterns allowed the dactyloscopists to limit the search for the source of an unidentified fingermark to a specific section of their databases of fingerprint reference files. Francis Galton proposed the first system of fingerprint classification in 1891, and the development and practical application of dactyloscopy for forensic use were materialized in 1892 with the publication of his manual of dactyloscopy. This led to the acceptance of fingerprints in Great-Britain and the British Empire. In 1900, Henry modified the classification system of Galton, which remained the most widely used system in the world under the name of Galton-Henry. In 1891, Vucetich began to collect the first ten print cards databases based on the ideas of Francis Galton and developed another classification system, which was adopted by some South-American countries. The size of the ten print cards databases increased progressively during the twentieth century, and the workability was maintained sophisticating the indexation system, but to the cost of a trade-off between selectivity and reliability. The coexistence of several classification systems around the world limited the interoperability of the manual classification between different systems. In the second part of the twentieth century, manual classification was slowly abandoned and replaced by computerized classification systems named Automatic Fingerprint Identification Systems (AFIS) [9].

Automatic Classification

Development

From the mid-1960s, research on automation of fingerprint identification started. USA and Japan concentrated on automation of the high-volume ten-print workload, while France and the UK focused more on automation of fingermark identification. After a decade of effort, digitization of the ten-print card and automatic designation of minutiae were effective enough for the USA and the UK to produce automatic fingerprint reader systems. This advancement opened the possibility to digitize the ten print card records and to store the standard impressions and the demographic data of individuals (e.g., name, citizenship, and date of birth) in a computerized database.

Forensic Uses of AFIS Technology

AFIS technology was initially developed to assist the dactyloscopists with computers in the identity verification process of individuals through their fingerprints. This process consists in searching the ten fingerprints of an individual in the database of standard impressions to verify if he or she is already present in the database and, if present, to check his or her demographic data. The AFIS technology has achieved enough maturity to ensure an identity verification process that is virtually error-free from the technological point of view, even if clerical mistakes in the database or in the running of the process can never be excluded.

In the 1990s the improvement of both AFIS and computer technologies allowed for the processing of fingermarks, exploited in two forensic processes. Fingermarks can be used for forensic investigation, in order to establish a link between a crime scene or an object and an individual. They can also be used for forensic intelligence to establish links between several crimes, even if the potential for links using marks depends on their limited quality.

In the 2000s the improvement of the computer mass-storage, in terms size and affordability, favored the constitution of large-scale palmprints databases. This development allowed for an extension of forensic investigation and forensic intelligence based on palmmarks. In most countries, the constitution of large scale palmprints databases is an ongoing process.

The challenge of standardization has only been solved recently, through the use of a common format, developed by the American National Institute for Standards and Technology (NIST), facilitating the computerized exchange of fingerprint and fingermark data between countries and agencies [10].

Individualization of Fingerprints and Fingermarks

History

The criminalist Edmond Locard enounced the first rule establishing a minimum number of minutiae

necessary for fingermark identification. During 1911–1912 he initiated the discussion of a numerical standard for the forensic identification of fingermarks, suggesting the following rule:

1. If more than 12 minutiae ("concurring points") are present, and the fingermark is sharp, then the identity is certain. The imperative requirement for the absence of significant differences is implicit.
2. With 8–12 concurring points, the case is borderline and the certainty of the identity depends on
   a. The sharpness of the fingermark.
   b. The rarity of the type.
   c. The presence of the core of the general pattern and the delta in the usable part of the mark.
   d. The presence of pores.
   e. The perfect and obvious similarity of the print and the mark regarding the width of the papillary ridges and valleys, the direction of the lines, and the angular value of the bifurcations.

   In these instances, the certainty of the identification can only be established following a discussion of the case by at least two competent and experienced specialists.
3. With less than 8 minutiae, the fingermark cannot provide certainty for the identification, but only a presumption proportional to the number of minutiae available and their clarity.

Principally the first two parts of this rule were largely adopted by the community of the dactyloscopists but, unfortunately, the third part of the rule remained largely ignored [5].

Current Practice

The current dactyloscopic practice has evolved from the body of knowledge developed about the fingerprint individuality and the forensic use of fingermarks. It is formalized in a 4-step procedure named ACEV (Analysis-Comparison-Evaluation-Verification).
This procedure consists in the analysis of the fingermark followed by the analysis of the fingerprint, the comparison of the fingermark and the fingerprint, the evaluation and the decision based on the observed similarities and discrepancies between the fingermark and the fingerprint, and the verification of the findings by a second dactyloscopist.

Despite the formalization of the identification procedure, the practice varies between continents and countries, and even within some countries. The evaluation step, in particular, is based either on a quantitative threshold or on a qualitative threshold.

Quantitative Threshold: Presence of a Numerical Standard

A majority of European and South American countries favor a purely quantitative approach for forensic individualization, by fixing a numerical standard and considering qualitative aspects such as the third level details as secondary. A formal identification is established only if a minimal number of corresponding minutiae between the observed mark and the fingerprint – and an absence of significant differences – is put in evidence.

The numerical standard differs between countries and sometimes also between agencies in the same country: Italy (16-17); UK (before 2000) (16); Belgium, France, Israel, Greece, Poland, Portugal, Romania, Slovenia, Spain, Turkey, South American Countries (12); Netherlands (10-12); Germany (8-12); Switzerland (before 2008) (8-12); and Russia (7) [5].

Qualitative Threshold: Absence of Numerical Standard

Until 1970, the fingerprint identification procedure in the USA was also based on a numerical standard of 12 points, and below this threshold, qualitative factors in the comparison were taken into consideration. In 1970, a commission of experts from the International Association for Identification (IAI) was established to study the question of the relevancy of a fixed numerical standard for dactyloscopy. The following resolution was adopted by the IAI in 1973: "The International Association for Identification, based upon a 3-year study by its Standardization Committee, hereby states that no valid basis exists for requiring a predetermined minimum of friction ridge characteristics that must be present in two impressions in order to establish positive identification."

It was accepted that the concept of identification could not be reduced to counting fingerprint minutiae, because each identification process represents a unique set of features available for comparison purposes; the identification value of concurring points between a fingerprint and a fingermark depends on a variety of conditions that automatically excludes any minimum standard.

In 1995, during a conference meeting on fingermark detection techniques and identification hosted in Ne'urim, Israel, 28 scientists active in the field of dactyloscopy, representing 11 countries, unanimously approved a resolution that is a slight variation of the IAI 1973 resolution. The Ne'urim declaration states that "no scientific basis exists for requiring that a predetermined minimum number of friction ridge features must be present in two impressions in order to establish a positive identification."

Decision Process

A formal identification is established when the dactyloscopists reach a decision threshold. They evaluate the contributions to individuality on a quantitative level (numerical standard), or on a qualitative level (absence of numerical standard), and the size of the relevant population of potential sources of the fingermark is set to its maximum, independently of the circumstances of the case [5].

On the basis of their evaluation, most dactyloscopists report three types of qualitative opinion: identification, exclusion, and inconclusive. As their evaluation is deterministic, they also make an implicit use of their own subjective probabilities of the rarity of the characteristics used to substantiate their opinion. They refine these subjective probabilities through training and experience, but they rarely consider results from research, particularly in the fields of embryology and statistics.

Admissibility of the Fingerprint in the USA

Like for other forensic disciplines, the scientific status of fingerprint identification has been questioned since 1993, when the Supreme Court of the USA handed down its ruling in Daubert v. Merrell Dow Pharmaceuticals (1993, Inc., 509 US, 579). Previously the main criterion for the admissibility of expert testimony in the federal courts of the USA was the Frye standard, which requires the general acceptance of the methods by the relevant scientific community. Daubert gave federal judges much greater discretion in deciding admissibility. It suggested that they consider (1) whether a theory or technique can be tested, (2) whether it has been subject to peer review, (3) whether standards exist for applying the technique, and (4) the technique's error rate. Although it is possible to test and validate methods for the forensic individualization of fingermarks, the research on this topic is still very limited.

The admissibility of fingerprint evidence, as being scientific in nature, has been subject to a Daubert hearing in the case U.S. v. Mitchell (1999, U.S. District Court for the Eastern District of Pennsylvania, Criminal), followed by Daubert hearings in more than 20 other fingermark cases. In the same case, U.S. v. Mitchell, the FBI provided calculations based on experiments carried out on an AFIS system. Random match probabilities of 10^{-97} and 10^{-27} were claimed respectively for complete fingerprints and partial fingermarks. These extraordinary numbers have been obtained by an extreme extrapolation of the probability density of the score using a postulated model, but they are so far from reality that it is surprising that they were admitted as evidence. Until January 2002, all Daubert hearings on fingermark cases led to the full admissibility of fingermark evidence in the courtroom. Judicial notice was given to the fact that fingerprints are permanent and unique [5].

January 2002 coincides with the first decision that proposes to limit expert testimony on fingerprint identification. Indeed in U.S. v. Llera Plaza (188F. Supp. 2d 549, 572–73 (E.D. Pa. 2002)), the defense "Motion to Preclude the United States from Introducing Latent Fingerprint Identification Evidence" has been partly successful. Judge Pollak held that a dactyloscopist could not give an opinion of identification, and required that the expert limits his testimony to outline the correspondences observed between the mark and the print, leaving to the court the assessment of the significance of these findings. That led the Government experts to ask for reconsideration bringing to the debate background documents in relation to the move of the UK toward the abandonment of the 16 point standard. Judge Pollak later reversed his opinion, and admitted the evidence.

Two cases of wrongful fingermark identification following the case of the Scottish police officer Shirley McKie perpetuated this controversy. In the first case the American Stephan Cowans was convicted by fingerprint identification, but later exonerated by DNA analysis. In the second case, the American Brandon Mayfield was wrongly associated with the 11 March 2003 Madrid bombing, by means of fingerprint to a latent mark revealed by the Spanish National Police on a plastic bag containing detonators recovered from a stolen van associated with these bombings. Three FBI experts and an independent court-appointed expert all identified Mayfield as the donor of the mark. Mayfield, a lawyer based in the US State of Oregon, came to the

FBI's attention when one of the latent marks sent by the Spanish authorities through Interpol gave a hit against his name on the FBI integrated AFIS (IAFIS), containing about 440 millions of fingerprints from 44 millions of persons. Brandon Mayfield was arrested, and remained in custody for a few weeks until the Spanish dactyloscopists, who immediately had raised issues with this identification, finally identified the mark with the finger of an Algerian suspect.

The FBI offered an apology and published a research report in the beginning of 2004 in which the existing FBI procedures were investigated extensively. This report showed that the mistake in this case was not owed to the methods the FBI used, but was the consequence of "human error" which cannot be excluded. The problem with this frequently used explanation is that the method and the human cannot be separated in case of an activity at which the human acts as a measuring instrument as is the case in traditional dactyloscopy [11].

An extensive research by the General Inspector of the US department of Justice appeared in January 2006 in which a clear analysis was given of the facts and circumstances causing the incorrect identification [12]. According to this report, an important factor in the Mayfield case was that when a search is performed using a very large database, there will always be a reference print which strongly looks like the unknown mark. A positive consequence of these cases is that they initiated a move towards a much more open discussion about the misidentifications in the forensic fingerprint field.

Analysis of the Current Practice

Research in embryology and statistics clearly do not legitimate the reduction of fingerprint individuality to counting minutiae. Indeed the scope of features is much broader than minutiae alone, and the nature of the papillary individuality prevents the adoption of any predefined number of ridge characteristics necessary for identification, without significant differences [13]. It is axiomatic that no two fingerprints are identical, as no two entities of any kind can be identical to each other. A common misconception lies in the fact that the features of individuality of the fingerprint is often attributed to the fingermark. As already described by Locard, in criminalistics, the transfer of material is logically never perfect. In dactyloscopy, the transfer of the pattern from the fingerprint ridges to the fingermark is accompanied by two types of loss of information: quantitative, due to the limited size of the trace, and qualitative, due to distortion, blurring, bad resolution, and loss of pore and edge details.

The challenge for dactyloscopy is about the ability to quantify the information available for the individualization process in a partial distorted fingermark, and not to prove the individuality of the friction ridge skin. The first step in the quantification of the evidential value of fingermark evidence consists in estimating the similarity between the features of this fingermark and those of the fingerprint considered as potential source of this mark. The second step consists in estimating the typicality or the rarity of these features, and the third step, in reporting the similarity–typicality ratio as evidential value. This concept encapsulates a continuum of values for individualization of the fingermarks ranging from very high to very low, depending on the feature analyzed. Therefore, the forensic individualization process of fingermarks cannot be considered as a binary decision process, but has to be envisaged as a purely probabilistic assessment of the value of evidence, as it is for any type of evidence [14].

Probabilistic models, which are applicable to fingermark individualization [15], have been proposed and accepted by forensic scientists in other forensic areas – i.e., DNA, microtraces and speaker recognition [16]. The absence of extensive statistical analysis on fingerprint variability can be viewed as the main reason to prevent giving qualified opinions. Statistical data only support and comfort identification statements used by dactyloscopists but, according to Stoney, "we must realize that to reach absolute identification, or its probabilistic equivalent, through an objective process is not possible. Probabilities are objective when they can be tested and reproduced" [17].

Future Perspectives

The statistical studies applied to fingerprints and fingermark individualization provide valuable knowledge about the statistical behavior of various types of features, mainly the minutiae, and to a more limited extent, the pores, but they do not provide a robust tool to assess the probability associated with a given configuration of features for several reasons: none of the proposed models has been subjected to an extended empirical validation, and the assumptions about

the features used in these models have not been fully explored.

The research possibilities are huge, mainly in three different directions. The first is a refinement and an empirical validation of the model-based approaches developed in earlier studies [8]. The second is the development of data-driven approaches taking advantage of the capabilities of the current AFIS systems, embedding large fingerprint and fingermark databases, high computation capabilities, and sophisticated pattern recognition techniques. The third direction is to explore the morphogenesis process from the point of view of mathematical biology, with the aim to determine the contribution of the genetic, environmental, and the other factors, which influence the features defined in the three levels of information present in the fingerprint. These studies require the availability of large samples of fingermarks and fingerprints and a clear definition of the features used by the examiners to compare fingermarks with fingerprints.

Related Entries

- ▶ Automatic Fingerprint Matching
- ▶ Fingerprint Classification
- ▶ Fingerprint Databases and Evaluation
- ▶ Fingerprint Features
- ▶ Fingerprint Individuality
- ▶ Fingerprint Matching, Manual
- ▶ Individuality of Fingerprints
- ▶ Latent Fingerprint Experts

References

1. Meuwly, D.: Forensic individualization from biometric data. Sci. Justice **46**(4), 205–213 (2006)
2. Kuecken, M., Newell, A.C.: Fingerprint formation. J. Theor. Biol. **235**, 71–83 (2005)
3. Wertheim, K., Maceo, A.: The critical stage of friction ridge and pattern formation. J. Forensic Ident. **52**(1), 35–85 (2002)
4. Jain, A.K., Prabahakar, S., Pankanti, S.: On the similarity of Identical Twin Fingerprints. Pattern Recognit. **35**(12), 2653–2663 (2002)
5. Champod, C., et al.: Fingerprints and other Ridge Skin impressions. CRC press, London (2004)
6. Ashbaugh, D.R.: Qualitative-quantitative friction ridge analysis – An introduction to basic and advanced ridgeology. In: Geberth, V.J. (ed.) Practical Aspects in Criminal and Forensic Investigations. CRC Press, Boca Raton, FL (1999)
7. Stoney, D.A.: Measurement of fingerprint individuality. In: Lee, H.C. Gaensslen, R.E. (eds.) Advances in Fingerprint Technology, pp. 327–388. CRC Press, Boca Raton, FL (2001)
8. Neumann, C., et al.: Computation of likelihood ratios in fingerprint identification for configurations of any number of minutiae. J. Forensic Sci. **52**(1), 54–64 (2007)
9. Berry, J., Stoney, D.A.: History and development of fingerprinting. In: Lee, H.C., Gaensslen, R.E. (eds.) Advances in Fingerprint Technology, pp. 1–40. CRC Press, Boca Raton, FL (2001)
10. McCabe, R. M. (ed.): Data format for the interchange of fingerprint, facial, scar mark & tatoo (SMT) Information, American National Standard ANSI/NIST-ILT 1-2000, July (2000)
11. Fine, G.E.: A Review of the FBI's handling of the Brandon Mayfield case. 2006, Office of the Inspector General, U.S. Department of Justice
12. Office of the Inspector General, United States Department of Justice. A Review of the FBIs Handling of the Brandon Mayfield Case: Unclassified Executive Summary, Washington, DC (2006)
13. Champod, C.: Dactyloscopy: Standards of proof, In: Siegel, J. (ed.) 1Encyclopedia of Forensic Science. Academic, London. (2000)
14. Taroni, F., Champod, C., Margot, P.: Forerunners of Bayesianism in early forensic science. Jurimetrics **38**, 183–200 (1998)
15. Good, I.J.: Weight of evidence and the Bayesian likelihood ratio, In: Aitken, C.G.G. (ed.) Statistics and the Evaluation of Evidence for Forensic Scientists. Wiley, Chichester, UK (1995)
16. Aitken, C.G.G., Taroni, F.: Statistics and the evaluation of evidence for forensic scientists. Wiley, Chichester, UK (2004)
17. Stoney, D.A.: What made us ever think we could individualize using statistics. J. Forensic Sci. Soc. **31**(2), 197–199 (1991)

Fingerprint, Palmprint, Handprint and Soleprint Sensor

Geppy Parziale
Cogent Systems, Inc., South Pasadena, CA, USA

Synonyms

Fingerprint device; Fingerprint sensor; Handprint sensor; Palmprint device; Palmprint sensor; Soleprint device; Soleprint sensor

Definition

A fingerprint or palmprint or handprint or soleprint sensor is a transducer that converts the ridge–valley structure of a person's hand or foot sole to an electrical signal. Generally, the sensor *reads* the difference of

pressure, temperature, light, electrical capacity or other kinds of energies are measured between the ridges and the valleys. Then, this difference is converted into an electrical digital signal that is encoded as an image representing the ridge–valley pattern. Different technologies can be applied to achieve this conversion and each of them brings advantages and disadvantages.

It is important to highlight that the output signal is a representation of the real-world ridge–valley pattern. Hence, if F is a ridge–valley pattern of a real-world finger tip and s is the transfer function of a device, the output signal is $F' = s(F)$ and $F' \neq F$.

Introduction

The similarity of the ridge–valley pattern of the epidermis present on finger tips, palms, and soles [1] allows to use the same physical principles for capturing fingerprints, palmprints and soleprints. The devices using these technologies can be grouped into a single family, known as ► livescan furrow devices or shortly, *livescan devices*.

The technological advancement of livescan devices has been mainly driven by the research done in the fingerprint recognition field more than the palmprint and soleprint modalities. The reason has to be found in a more convenient use of fingerprint devices, instead of the larger, heavier, and more power-consumer palmprint and soleprint ones. Moreover, fingerprint being the oldest biometric means used to identify people, large collection of data have always been available. This facilitated the development of algorithms for fingerprint recognition, pushing experts and scientists to focus mainly on this modality more than palmprints and soleprints.

Ink-on-Paper Method

The oldest approach to capture the furrow pattern is represented by the ink-on-paper method. Even if we cannot consider this as a real sensing technology, it is important to mention it here, since ink-on-paper is still widely used to collect palmprints, fingerprints, handprints, and soleprints. Moreover, it represents a strong obstacle for the advancement and the introduction of new capture technologies. The reason has to be found in the existence of very large databases collected using this method during the last ten decades. When a new technology is introduced on the market, it must have a high degree of interoperability with the ink-on-paper method to ensure the continuity of the use of these databases, because a fingerprint or palmprint representation different than the legacy one would make the comparison very difficult. Thus, the representation of the ridge–valley pattern provided by the ink-on-paper method still represents the *model* that the modern technology tries to imitate.

The ink-on-paper capture approach consists in covering the ridge–valley pattern with black ink. Then, the print is obtained impressing the inked skin onto a white paper applying a small pressure. The resulting print is represented by a black mark for each ridge, while nothing is left in correspondence of each valley. The quantity of the ink applied on the skin and the pressure applied onto the paper during the impression are very important factors influencing the quality of the final result. In spite of other approaches, this technique does not suffer the skin condition problems (dry skin, wet skin, etc.), which are instead very difficult to overcome in the case of the other capture methods.

In some applications, the capture of fingerprints is performed rolling the finger onto the paper. This is done to acquire as much information as possible of the finger tip that can be used during an identification. The impression obtained with this approach is called ► rolled-equivalent fingerprint. Using dedicated image processing algorithms, rolled-equivalent fingerprints can also be obtained rolling the finger on the sensing surface of a sensor.

Collecting fingerprints, palmprints, handprints, and soleprints with the ink-on-paper method is still widely used, because it still represents the cheapest way to collect these biometric data. In Spain, the registration of all new born children is done applying ink on the baby soles of the feet and impressing them on a paper. In some Asian countries, inked fingerprints are used to register civilians during elections to avoid double voting. In Switzerland, Spain, Germany, and many other Countries, criminals are still registered by inking the tips of their fingers and their hands to collect the ten flat and rolled fingerprints and palms.

Once the fingerprints are collected on the paper, they can be digitalized using a flatbed scanner and then stored in digital format. This approach is still the most used by an Automated Fingerprint Identification System (AFIS) or an Automated Palmprint and Fingerprint Identification System (APFIS).

To improve the user convenience, especially in civilian applications, a special transparent oily substance is used in place of the black ink. In this way, the fingerprinted person does not need to wash her/his hands many times to remove the inconvenient residues of the black ink.

Nowadays, palmprints are becoming more and more popular in crime investigation, especially for latent comparison, since recent studies have demonstrated that more than 30% of the latent prints found on a crime scene belongs more likely to palms than fingers. Generally, the ink-on-paper palm capture consisted of inking the lower palm and the impress it on a paper. Nowadays, the reduction of the cost of the digital storage space allows to store larger quantity of data. Thus, the most modern approach consists of capturing the full handprint consisting of the three ▶ palm segments (lower, upper and writer palms).

Sensor Characteristics

Before describing the modern fingerprint sensor technologies, their main characteristics are highlighted in this article. These features define the application range in which the sensor can be used. For some applications, livescan devices have to pass very strict tests. The most famous and required certification is the *FBI fingerprint scanner certification*, covered in the Appendix F of the Criminal Justice Information Service (CJIS) Electronic Fingerprint Transmission Specification [2]. A list of FBI certified livescan devices is available at http://www.fbi.gov/hq/cjisd/iafis/cert.htm.

The first important feature for a livescan device is the *Image Resolution*, which describes the ability of a sensor to distinguish, detect, and/or record physical details of the ridge–valley pattern. It represents the number of pixels in a unitary length and is expressed in *pixel-per-inch* or shortly, ppi. Typical image resolution values are 500 and 1,000 ppi. The first value is the most common and it is used in majority of the applications and products present in the market. The 1,000 ppi is mainly used for criminal investigation, especially for palmprints. The interest in analyzing the so-called third level details of a ridge–valley pattern is now pushing the manufacturers to introduce new devices with very high resolution (1,500–5,000 ppi).

Radiometric Resolution or *Image Depth* or *Dynamic Range* determines how finely a sensor can represent or distinguish differences of intensity. It is usually expressed as a number of gray levels or bits, for example, 8 bits or 256 gray levels which is typical of fingerprint image.

The *Modulation Transfer Function* (MTF) or *Spatial Frequency Response* is another important parameter. Spatial frequency is typically measured in cycles or line pairs per millimeter (*lp/mm*). The more extended the response, the finer the detail and the sharper the image. MTF is the contrast at a given spatial frequency f relative to contrast at low frequencies and it can be computed with the following Eq. (1):

$$MTF = 100\% \frac{C(f)}{C(0)}, \tag{1}$$

where $C(f) = (V_{max} - V_{min})/(V_{max} + V_{min})$ is the contrast at frequency f, and $C(0) = (V_W - V_B)/(V_W + V_B)$ is the low frequency contrast. V_B, V_W, V_{min} and V_{max} represent the luminance for black areas, the luminance for white areas, the minimum luminance for a pattern near spatial frequency f and the maximum luminance for a pattern near spatial frequency f respectively.

All the optical features of a sensor can be measured using special targets. To test the quality of a device, a manufacturer must purchase these targets and test the accuracy of all its optical features.

The *Geometric Image Accuracy* represents the absolute value of the difference $D = X - Y$, between the distance X measured between any two points on the target and the distance Y measured between those same two points on the output image. This is a very important parameter especially for devices having a very large capture area.

The *Grayscale Linearity* (GL) represents the capacity of a device to reproduce the gray level values correctly. A target with gradually varying grayscale levels is used for this scope. The grayscale levels on the output image are compared with the grayscale levels on the input target to measure the accuracy of the representation.

The *Signal-to-Noise Ratio* (SNR) is measured using another special target representing a grayscale level reference. This reference can be a white-colored and a black-colored target. An image is generated from these two targets and compared point by point with the reference.

The *Framerate* represents the number of frames a sensor can generate per time unit. It is measured in *frames/s* and it is a very important parameter when the

object (finger, palm or hand) movements are implied during a capture (sweep devices). To improve the final image quality, many sensors acquire more images of the same finger or palm during an acquisition. The captured images are then combined to produce the final image.

The *Shutter-speed* is the time that a detector needs to capture a single image.

Other important sensor characteristics that can change the application range of a device are the communication interface type (USB, Firewire, Ethernet, etc.), the sensor dimensions and weight, the Mean-Time-Before-Failure (MTBF), the self-powering capacity (if the power of the device comes from the communication interface) and obviously the price.

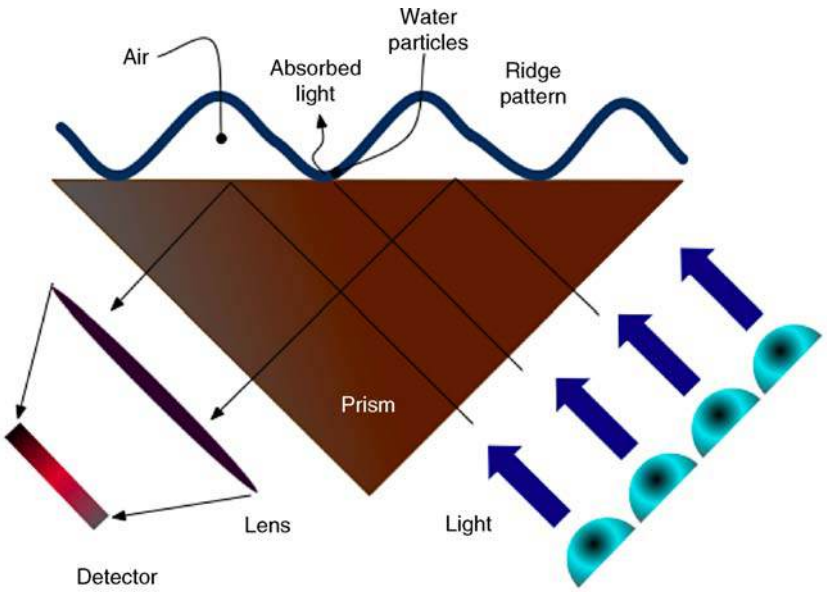
Optical Sensors

Sensors using light to discriminate between ridges and valleys represent the oldest technology to capture fingerprints, palmprints, and soleprints with no need of inking hands and feet.

The most widely used optical capture principle is known as Frustrated Total Internal Reflection (FTIR) and highlighted in Fig. 1. The sensor contains an optical prism and one of its faces is used as the ▶ platen that must be touched by the finger to produce an image. A monochromatic light source enters the prism and is reflected in accordance with each valley of the skin. Then, it is collected by a detector (CMOS or CCD array). The light is absorbed in accordance with each ridge touching the platen. The lack of reflection allows the ridges (appearing dark in the image) to be discriminated by the valleys (appearing bright in the image). The 3D ridge–valley structure plays here an important role: presenting to the platen a picture or a drawing of a fingerprint does not produce any image. On the other hand, molding the shape of the ridge–valley pattern with special materials (latex, silicon, etc.) and touching the platen with it produces an image that cannot be distinguished by the image obtained by the real ridge–valley pattern (spoofing).

This capture technology is strongly influenced by the skin conditions. When the skin is too dry, the ridges do not completely adhere to the glass platen and thus, an image with very low contrast is obtained. On the other hand, very wet fingers produce an uniform black spot image, because during the finger pressure, the sweat accumulates in accordance with each valley and the light of the LEDs is fully absorbed. The result is a uniform dark spot with very low contrast between ridges and valleys. To overcome these problems, before each capture the user is asked to clean her/his hands and wet them with special non-toxic substance providing the right quantity of wet on the skin.

Another problem related to optical devices is represented by the so called ▶ halo effect. When the wet skin touches the colder platen, the moisture starts to



Fingerprint, Palmprint, Handprint and Soleprint Sensor. **Figure 1** Total frustrated internal reflection (▶ TFIR) principle.

condense. This results in a halo on the final image reducing the ridge–valley contrast. To avoid the halo effect, the platen has to be warmed up. This process is expensive in terms of current consumption especially for palmprint and soleprint devices with very large platen. The warming process requires a certain period of time and thus, these devices cannot be used immediately after they are switched on. This is sometimes impractical for some applications.

The physical size is another limitation of the optical sensors. The length of the optical path (the path traversed by the light from its source to the detector) cannot be significantly reduced without introducing severe optical distortions on the final image. The use of small mirrors and special lenses can help in keeping the same path length in a small space, but the manufacturing costs drastically increase and the robustness of the device decreases.

Nowadays, optical sensors represent the maturest technology in the market for capturing fingerprints, palms and soleprints. The large production of this kind of devices is reducing their price more and more. Since they are rugged and less sensitive to environmental factors than other technologies, optical sensors are spreading very fast and blocking the penetration of other capture technologies into the market.

Palmprint devices, ► slap or four-four-two devices and ► Rolls Capture Devices are only available based on this technology for the quality it can provide also in the case of devices with large platens. This is why palmprint sensors are only available based on optical technology.

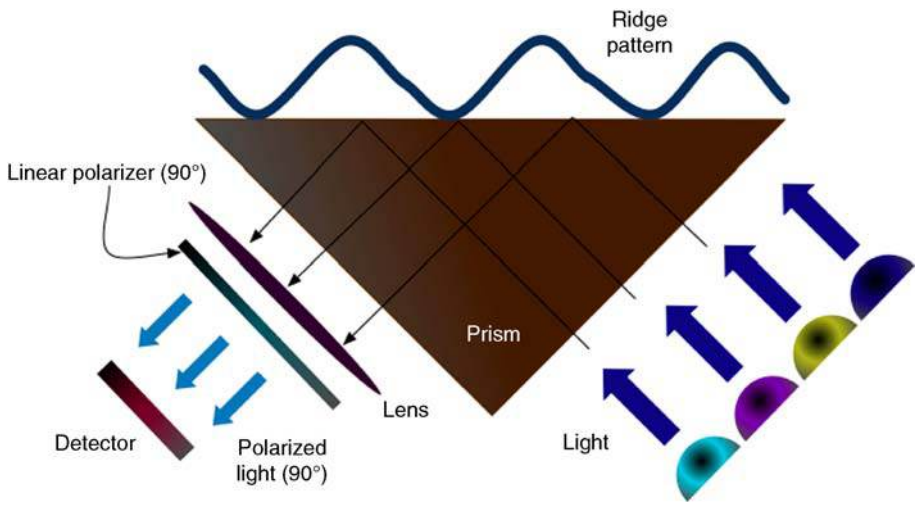
Optical Multispectral Sensors

To improve the ridge detail representation provided by the FTIR method, a novel approach to capture fingerprint has been recently proposed [3–5]. This approach is called Multispectral Imaging (Fig. 2) and uses multiple illumination wavelengths rather than a single monochromatic illumination commonly used in the FTIR approach. The orthogonal configuration of linear polarizers emphasizes this multispectral light, which penetrates the surface of the skin. The light then undergoes multiple scattering events before emerging from the skin toward the image array. In avoiding the optical phenomenon of the FTIR, the multispectral imaging sensor is capable of collecting more identifying data from the finger than the FTIR sensor. Currently, only fingerprint devices are available based on this technology.

Optical Multispectral Imaging is also claimed to be capable of detecting fake fingers obtained with organic or synthetic materials. The difference between the spectral characteristics of the skin and these materials is known and can be used to detect fake fingerprint.

Optical Contactless or Touchless Sensors

When a finger touches or rolls onto a surface, the elastic skin deforms. The quantity and direction of the pressure applied by the user, the skin conditions, and the projection of an irregular 3D object (the finger) onto a



Fingerprint, Palmprint, Handprint and Soleprint Sensor. **Figure 2** Multispectral imaging principle.

2D flat plane introduce distortions, noise and inconsistencies on the captured fingerprint image. To overcome these problems, a new approach to capture fingerprints has been proposed [6, 7], called touchless or ►contactless *fingerprinting*. Because of a lack of contact between the finger and any rigid surface, the skin does not deform during the capture and the repeatability of the measure is improved.

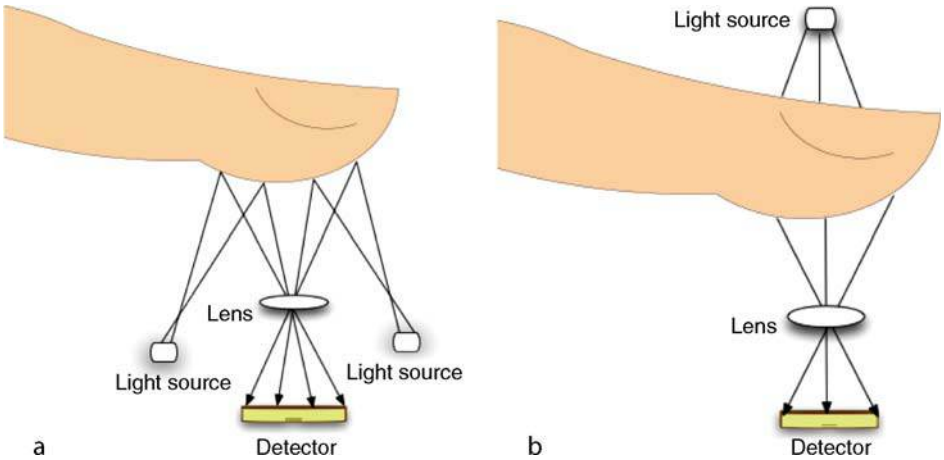
The approaches used to capture a fingerprint based on touchless technology can be grouped in two main families: ►Reflection-based Touchless Finger Imaging (RTFI) and ►Transmission-based Touchless Finger Imaging (TTFI). Figure 3 highlights the two approaches. In the RTFI approach, the light generated by monochromatic light sources and reflected on the finger skin is collected by the detector. In the TTFI approach, the light penetrating the finger is collected by the detector positioned in front of the ridge–valley pattern.

Since both the light reflecting on or penetrating the valleys and the light reflecting on or penetrating the ridges are collected by the detector, the final image has a contrast lower than that in the traditional FTIR technology. This has a huge impact on the minutiae extraction algorithm and thus, the advantage of a lack of skin deformation is negatively compensated by this low contrast. Moreover, the illumination not being perfectly perpendicular to the skin surface, shadowing effects of the ridges on the valley provide a wrong representation of small details (minutiae, pores, island, branches, etc.). Sophisticated illumination techniques are required to avoid this representation problem and increase the final image contrast. The consequence is an increase of the size and final costs of these devices.

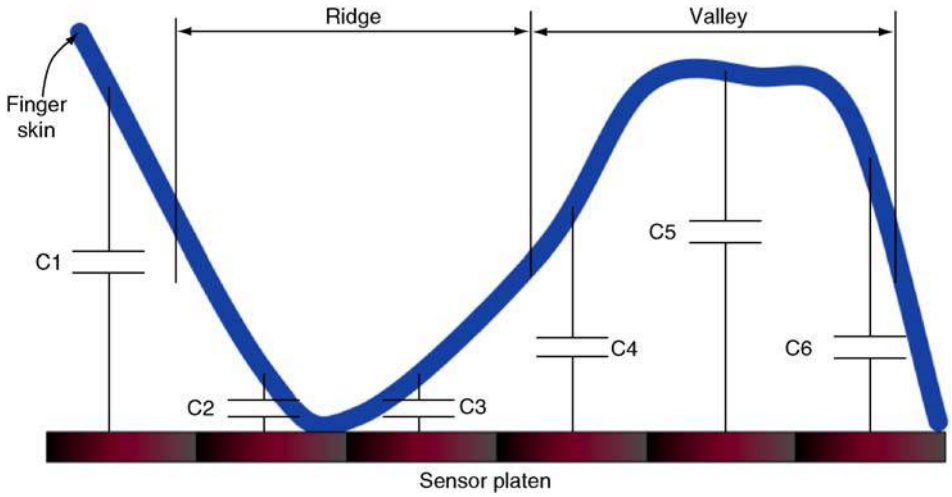
Another disadvantage of this technology is represented by the easy methods that can be used to attack these devices, which cannot be definitively used for high-security applications. In contrst to the FTIR case, where the ridge–valley 3D structure is important to generate an image, the touchless approach cannot discriminate between a 2D and a 3D pattern. Hence, presenting a photograph or a simple drawing of a fingerprint to the sensor, a new fingerprint image similar to the synthetic one is generated and the access is granted. Finger positioning, sensor usability, and user convenience must be still addressed.

Solid-State Sensors

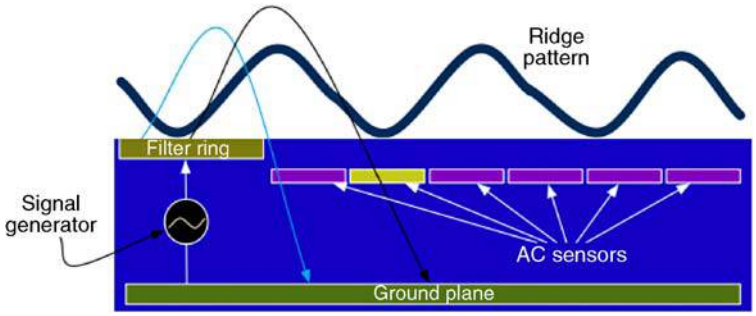
The first solid-state fingerprint capture device appeared on the market only in the middle of 1990s. It was a CMOS sensor capable of measuring the electrical capacity between the finger skin and the sensing surface (Fig. 4), which is composed by many squared pixels. Each pixel and the corresponding skin portion can be considered as an electrical capacitor with capacity $C = \varepsilon A/d$, where A represents the pixel area and d the distance between the skin and the pixel and ε is the permittivity (a constant depending on the material) of the dielectric contained between the two capacitor



Fingerprint, Palmprint, Handprint and Soleprint Sensor. Figure 3 ►Touchless or contactless capture approach: (**a**) reflection-based touchless finger imaging; (**b**) transmission-based touchless finger imaging.



Fingerprint, Palmprint, Handprint and Soleprint Sensor. **Figure 4** Capacitive principle for the capture of the ridge–valley structure.



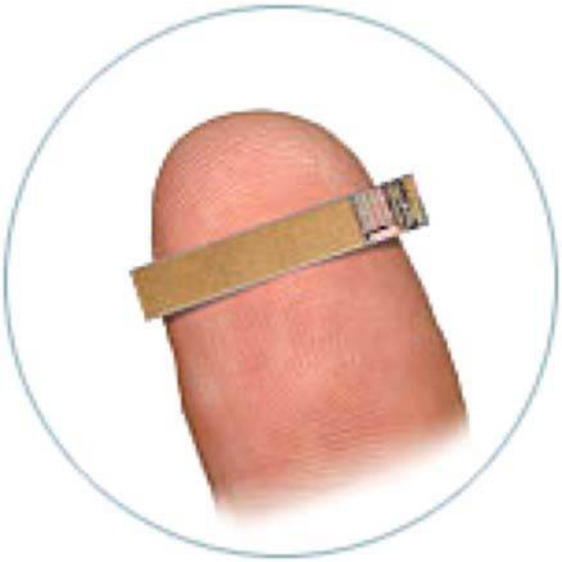
Fingerprint, Palmprint, Handprint and Soleprint Sensor. **Figure 5** Radio Frequency Field principle used to capture the ridge–valley structure.

plates. Each pixel produces a graylevel value proportional to its distance from the skin.

Another approach used to capture the ridge–valley pattern is based on the Radio Frequency (RF) electrical field (Fig. 5). A signal generator produces a low-level RF field traveling through the finger. The signal is then collected by AC sensors after being attenuated by the finger skin. The attenuation level of the signal is a function of the ridges and the valleys; the sensor array calculates the attenuation to synthesize the fingerprint structure. RF signal can be dynamically optimized in frequency and level to obtain the best possible image.

Using pyroelectric materials, it is possible to measure the difference of temperature between ridges and valleys. This approach is used while the finger is swiped on the small sensor surface (Fig. 6). This type of devices are called ► sweep sensors [8, 9]. The thermal sensing elements detect temperature difference between valleys and ridges during the finger movement. This technology is claimed to overcome the skin condition issues of optical sensors. However, the resulting images are not rich in gray level values, i.e., dynamic range. Sweep sensors are very attractive because of their small size and low cost. This makes easier their integration in handheld and mobile devices.

The big advantage of the solid-state technology is represented by their smaller dimensions and lower costs with respect to the optical technology. Since they can be manufactured very thin and their power consumption needs are very low, solid-state sensors can be mounted on cards, handheld devices or laptops and used as logon means. This has an implication on the range of applications in which solid-state fingerprint sensors can be involved with respect to the



Fingerprint, Palmprint, Handprint and Soleprint Sensor. Figure 6 An example of sweep sensor.

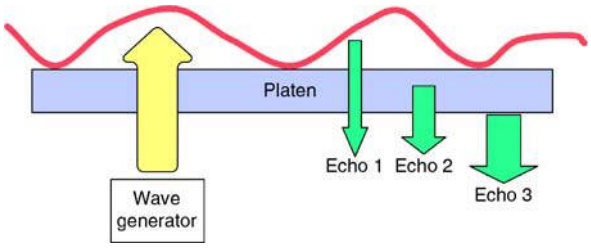
optical devices. However, external environmental factors (temperature, humidity, dust, etc.) are the major drawbacks of this technology. The sensing area is a chip completely open to the external world. Thus, special electrostatic protection methods must be used to avoid that external electrostatic charges destroy the chip surface. The same human skin can be the cause of the surface destruction, since the human body is usually electrically charged. Dust is another common vehicle of electrostatic charges that can quickly and easily degrade the sensing surface characteristics.

The use of solid-state sensor is mainly limited to fingerprint for their small sensing area. Palmprint and soleprint would require very large silicon areas that would make these sensors completely unaffordable in term of costs.

Even if their introduction on the market has been revolutionary and the expert envision new kinds of applications for fingerprint recognition (domotic, health-care, id-card and credit-card protection, etc.) their reduced lifetime and their high sensitivity to the external environmental factors limit the wide-spreading of these devices.

Ultrasonic Sensors

The ability to obtain images using ultrasound is based upon the reflection and transmission coefficients of



Fingerprint, Palmprint, Handprint and Soleprint Sensor. Figure 7 Capture principle of an ultrasonic capture device.

ultrasound as it propagates through media of varying acoustic impedance. What makes sound waves valuable for the imaging of the ridge–valley pattern is that they can both reflect and pass through objects. The characteristics of sound waves make it possible for high-frequencies to pass through substances and accurately measure the ridges and valleys of a fingerprint even if in presence of dirt, grease, ink, moisture, dye, or other substances routinely found on fingers.

The capture principle of a ultrasonic device is highlighted in Fig. 7. An ultrasonic wave generator produces high-frequency sound impulses. These impulses reflect on each material found on their path producing echos. The strength of each echo depends on the material and the shape of the object on which they were generated. Special receptors are used to translate the echos in an electrical signal.

Livescan imaging the fingerprints of children 5 years and younger is a technically challenging task, since the ridge structure is usually very fine and contains high "spatial frequencies," meaning that the ridges very close together. The spatial frequency of the fingerprint directly determines the resolution that the imaging device needs to accurately image the finger. Most live-scan fingerprint scanners have been designed to image adult fingers where a high-resolution scan is unnecessary. High-resolution ultrasonics is the only technology that can reliably and repeatedly capture clear and useful images of a young child's fingerprint.

Next Generation

Although most of the technologies mentioned earlier are quite new (some of them are still in the prototyping phase), the research and the development continues to

bring new ideas to this field. The study of the physiology and the formation of the furrow pattern allowed to propose new fingerprint and palmprint capture approaches. It is important to mention here the Optical Coherence Tomography (OCT) which is an interferometric, noninvasive, optical tomographic imaging technique offering millimeter penetration (approximately 2–3 mm in tissue) with micrometer-scale axial and lateral resolution. OCT is like an optical version of ultrasound imaging. The technique is already routinely used in medicine, but has not had a forensic application until now. The technique provides a transparent 3D structural picture by sending light though the pattern of natural secretions left on a surface by a finger and combining the reflected beam with a "reference beam" produced by bouncing light from a laser off a mirror. This produces an interference pattern at a photodetector the same as those found in a digital camera which can then be used to reconstruct an image of the original fingerprint.

This technology together with multispectral and touchless imaging must be still further developed to demonstrate their superiority with respect to the FTIR approach that still remains the most used method to capture fingerprints and palmprints.

Summary

Livescan furrow sensors represent a family of devices used to capture fingerprints, palmprints, handprints, and soleprints. The same anatomical characteristics of the skin present on finger tips, palms, and soles allow the use of the same technology for the capture of these biometric treats.

The ink-on-paper method is first method used to capture fingerprints and palmprints. Optical devices try to overcome the inconvenience of the ink on the skin and provide a good alternative method to the legacy ink-on-paper. Solid-state sensors are very attractive for their very small size and reduced costs, but they can only be used to capture fingerprints. Moreover, environmental factors limit the life time of these devices.

Multispectral and touchless imaging technologies try to overcome the limitation of the optical devices, but their relatively higher costs and very low interoperability with legacy technology limit their wide-spreading.

Related Entries

- ▶ Biometric Recognition
- ▶ Biometric Sensor and Device, Overview
- ▶ Fingerprint Recognition, Overview
- ▶ Fingerprint Verification

References

1. Ashbaugh, D.R.: Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology. CRC, Boca Raton (1999)
2. Criminal Justice Information Services Division: Electronic Fingerprint Transmission Specification. Department of Justice (1999)
3. Rowe, R.K., Nixon, K.A.: Fingerprint enhancement using a multispectral sensor. In: Proceedings of SPIE Conference on Biometric Technology for Human Identification, vol. 5779. pp. 81–93. Orlando, USA (2005)
4. Rowe, R.K., Corcoran, S.P., Nixon, K.A., Ostrom, R.E.: Multispectral Imaging for Biometrics. In: Proceedings of SPIE Conference on Spectral Imaging: Instrumentation, Applications, and Analysis, vol. 5694, pp. 90–99. Orlando, USA (2005)
5. Rowe R.K., Nixon, K.A., Butler, P.W.: Advances in Biometrics. Sensors, Algorithms and Systems, chap. Multispectral Fingerprint Image Acquisition. Springer, Berlin (2008)
6. Parziale, G.: Advances in Biometrics. Sensors, Algorithms and Systems, chap. Touchless Fingerprint Technology. Springer, Berlin, New York, USA (2008)
7. Parziale, G., Diaz-Santana, E.: The surround imager: a multicamera touchless devide to acquire 3D rolled-equivalent fingerprints. In: Proceedings of IAPR International Conference on Biometrics (ICB), pp. 244–250. Hong Kong, China (2006)
8. Parziale, G., Bishof, H.: Image reconstruction and on-the-fly minutiae extraction of fingerprints acquired with sweep sensors. In: Proc. of 28th Workshop of the Austrian Association of Pattern Recognition, pp. 241–248. Austria (2004)
9. Clausen, S.: Advances in Biometrics Sensors, Algorithms and Systems, chap. A single-line AC capacitive fingerprint swipe sensor. Springer, New York, USA (2008)

Fingerprints Hashing

Jean-François Mainguet
Grenoble, France

Synonyms

Biometric encryption; Cancelable biometrics; Fingerprint encryption; Fuzzy extractor; Fuzzy vault; Intricated biometrics

Definition

Fingerprint hashing is merging fingerprint recognition and cryptographic methods. The aim is to perform a recognition using fingerprint while, at the same time, hiding the private information related to the fingerprint, thus enabling public fingerprint templates.

Introduction

Keeping a database in a safe place is not easy. Even with good encryption methods and special care, databases containing sensitive information, such as bank account numbers, are vulnerable to being compromised. Nobody wants something like that to happen when dealing with fingerprint identification.

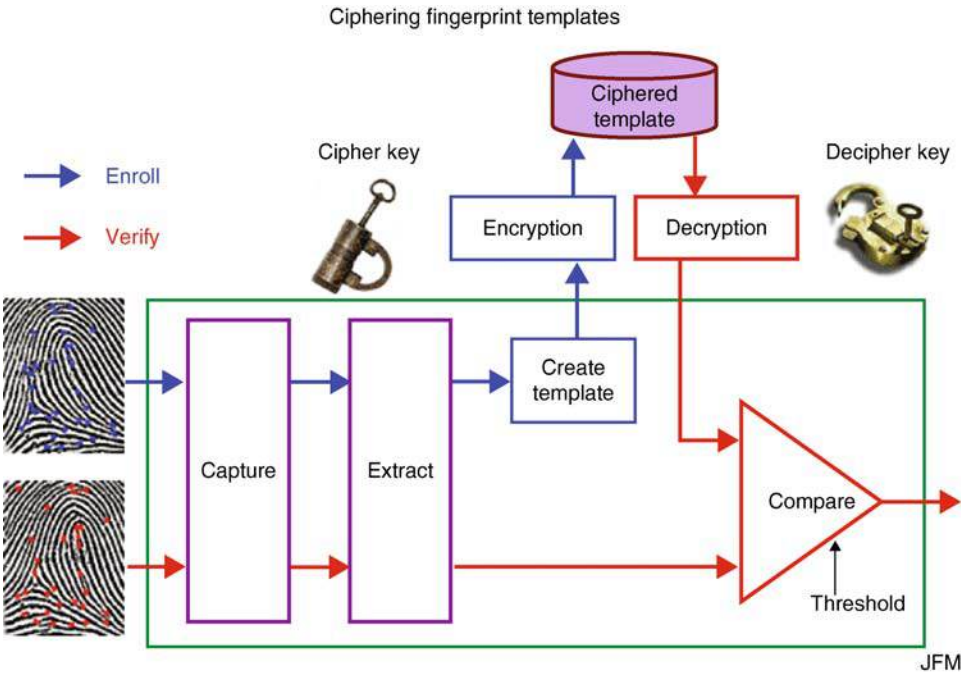
Security of a fingerprint-based system can be divided into two main areas:

1. The electronic security, which poses the question: "Is the electronic system, at the other end of the wires, a real trustful authorized fingerprint system?"
2. The liveness security, which asks a different question: "Is the object touching the sensor a real finger, alive and connected to a living person?"

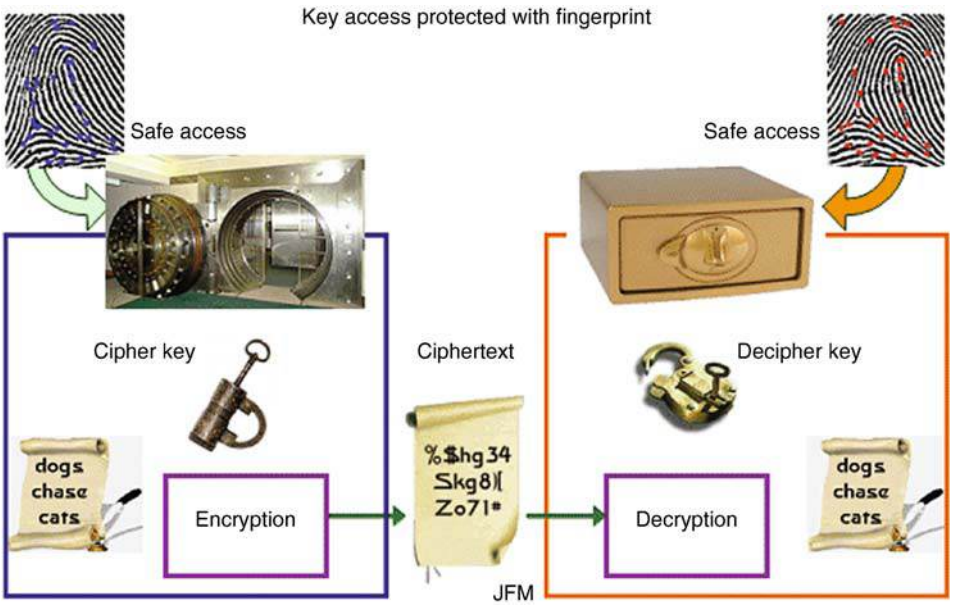
Liveness security is not addressed in this essay. Fingerprint hashing is part of the electronic security solution and deals with encryption.

Any biometric system requires the storage of a template (or reference). For fingerprint systems, the most common method consists in storing the minutiae. This information is considered as private information and should be protected and ciphered, not only for privacy reasons, but also against template replacement. This prevents a hacker from replacing the owner's minutiae, or using reverse engineering to get the minutiae locations and create a fake fingerprint. Even if biometric data cannot be considered as secret (they are public information in its cryptographic sense, always hiding ones face or voice is impossible), it is important to protect biometric data and the additional data that goes with them (name, bank account or whatever).

The template storage problem is generally solved using encryption. A template database can be created and protected using a single key pair (Fig. 1). Everything is fine up to the date when the key is compromised or badly protected, allowing hackers to access the data. However, in the case of fingerprints, this is ones very private information that is stored, and so it seems desirable to have an even better security scheme.



Fingerprints Hashing. Figure 1 Protecting fingerprint templates: Key is not protected.



Fingerprints Hashing. Figure 2 Protecting access to cryptographic secret keys using a fingerprint system: Fingerprint template is not protected.

Another problem occurs when one tryies to create a system enabling encryption/decryption features (like ► Pretty Good Privacy (PGP)). One needs to protect the access to the encryption/decryption module, which is done using a password. If one tries to replace the password with a fingerprint (Fig. 2), then one faces the problem of protecting the template, and cannot use the encryption/decryption scheme, because it is not yet enabled! It is the same problem as "you cannot put the key of the safe inside the safe itself." One needs another safe.

There is also an additional problem from a security point of view. The result of matching is only one bit of information that is easy to find and hack (too low entropy). It would be better to eliminate this weakness.

Desirable Features, Definitions

A better fingerprint system includes:

1. The storage of the template (minutiae) in a nonreversible way. It is still possible to perform a match, but it is impossible to recover the original minutiae and impossible to derive the secret key.
2. It is possible to revoke (cancel) a template. If a template is not to be used anymore, it is possible to forbid its use and create a new one.
3. There is no step with a single yes/no bit corresponding to the match/no match result.

Properties #1 and #2 are generally linked, because it would be very impractical and dangerous to use a transform that is unique. Each individual would have a unique number ID for his or her whole life, impossible to change.

Fingerprint hashing is the use of a non-reversible transform (similar to a hash function) over a fingerprint. It is also called "cancellable biometrics," because it is possible to cancel or to revoke the template. Fingerprint hashing involves using some kind of cryptographic scheme, similar to a hash function, but it is not a hash function.

Property #3 requires a stronger merge between biometrics and cryptography. Having all the properties at the same time is pretty hard to achieve and to prove, but has been originally proposed under the name of "Biometric Encryption" [1, 2]. Unfortunately, "biometric encryption" can be a simple combination of a biometric template and a simple encryption scheme. But it is much more; it is a real merge. In quantum cryptography, the word "► intricated" is used to designate the non-separable nature of some properties in quantum mechanics, and so "Intricated Biometrics" seems a better designation.

Cancelable Biometrics

Fingerprint hashing seems pretty close to password protection. A password is protected using a hash function, which is basically a method to transform some data into a relatively small number, the hash value, sometimes called fingerprint (which causes confusion), because of its uniqueness property (no collision should occur). A hash function is not reversible, and in most cases, some original data is lost as the resulting hash value is much shorter. This works well for password storage. You just need to apply the same hash function to the proposed password and perform a bit-to-bit comparison for checking. It is not useful to regain access to the original password.

Unfortunately, this scheme cannot apply to a fingerprint, because you never enter exactly the same fingerprint image. Every acquisition is different, and usual hash functions will return a different value, forbidding a further comparison. The problem is much more complex as it needs to be accepted that there will be some variability of the data. Fingerprint hashing must use a non-reversible transform like a hash function, but the comparison stops here. The other properties of a hash function, such as fixed length and uniqueness, are not required, but there needs to be a comparison, a match at the end, as depicted in Fig. 3.

General concepts related to cancelable biometrics have been discussed by Ratha et al. [3]. Davida et al. [4] added data to create a non-reversible template. Linnartz and Tuyls [5] proposed the use of specific shield functions before a hash function.

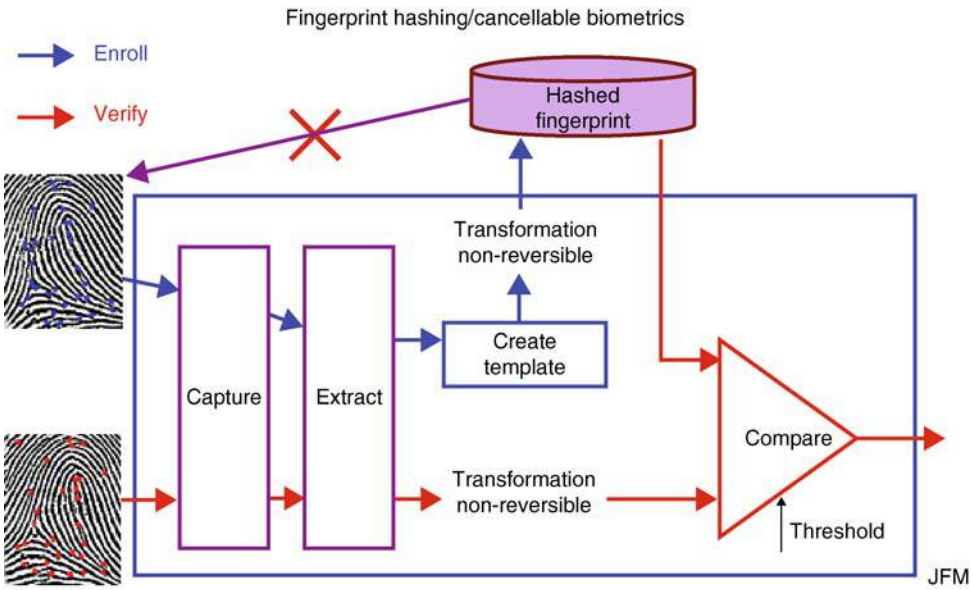
Lumini and Nanni [6] proposed the BioHash, a combination of a hash code, a Gram–Schmidt normalization and using a Hamming distance for comparison. This has been tested using the FVC-2002 fingerprint database.

Boult et al. [7] proposed another scheme called BioToken, and also tested on the FVC-2002 and 2004 fingerprint databases, which showed some enhancements of the accuracy of the system.

As usual in cryptography, proving that the transform is non-reversible or reversible with an extremely long computation time is very hard to achieve. Although there are some good reasons to believe that some solutions exhibit the right properties, nothing is mathematically proven yet. It took a long time for the security of regular cryptographic schemes to be accepted, and biometrics is in a similar situation, still in its infancy.

Intricated Biometrics

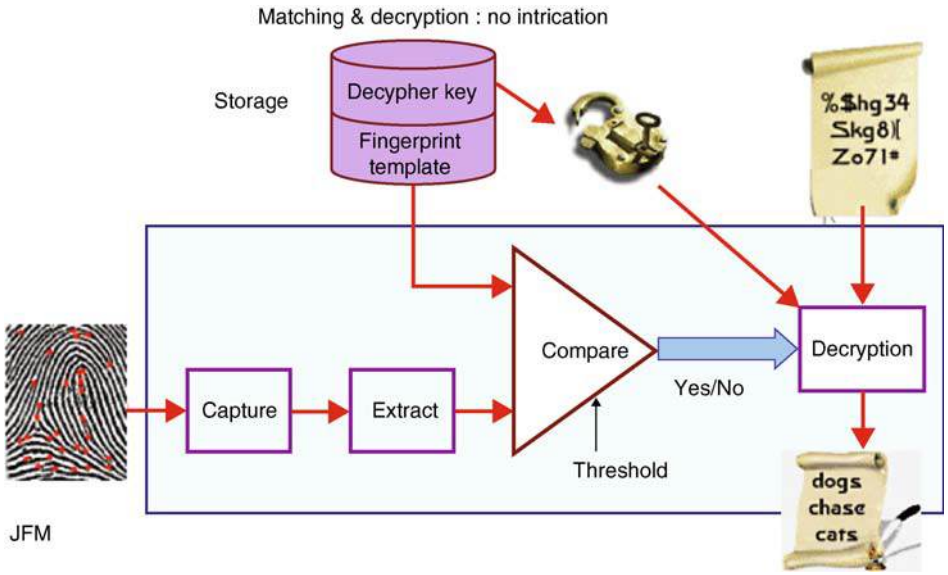
Cancelable biometrics shows interesting features, but still shows the potential weakness of the



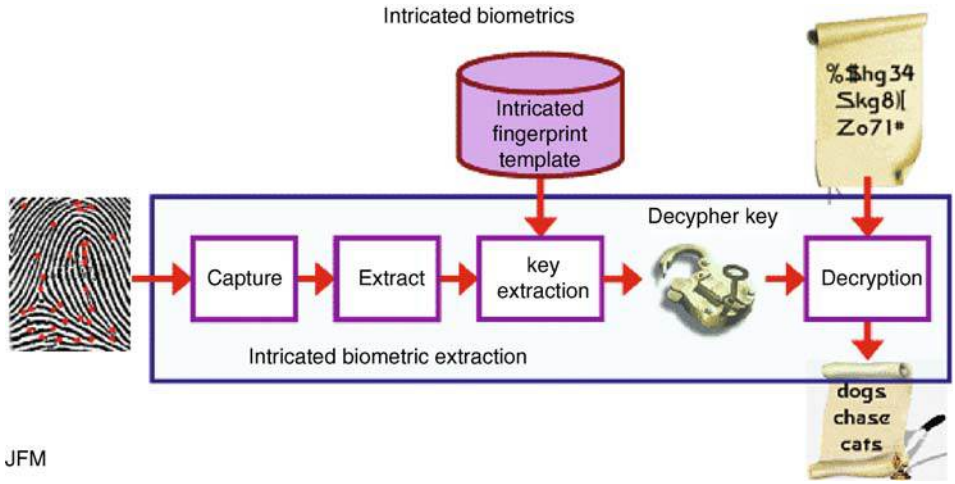
Fingerprints Hashing. Figure 3 Fingerprints hashing/cancellable biometrics: It is not possible to extract the original biometric data from the template as the transform is non-reversible.

match/no-match bit. This type of weakness may not be as critical when speaking of a physical or logical access, because of the need of a go/no go answer. But in most cases, the aim of the biometrically enabled system is to provide a service, and secure systems always use somewhere a cryptographic key when a transmission is involved in a non-secured environment.

A simple scheme uses the result of the match to enable the decipher key, as shown in Fig. 4. The template and the key are not protected. This requires external means; another secret key and method. There is still the one bit match/no match result. Intricated biometrics proposes to merge the decipher key with the template, so that neither the biometric template (the minutiae) nor the decipher key can be obtained from the stored template alone; they are intricated (Fig. 5). The intricated biometric template can be stored anywhere, even in a non-secured area.



Fingerprints Hashing. Figure 4 A simple use of biometrics to use a secret key: fingerprint and secret key are not yet protected, and the 1-bit match/no match still exists.



Fingerprints Hashing. Figure 5 Intricated biometrics: It is not possible to get the secret key and the original fingerprint data from the intricated fingerprint template, and the extracted key appears only for a short while for deciphering. The 1-bit match/no match step is eliminated.

When the decipher key is to be used, the live fingerprint can be scanned. If the extracted minutiae corresponds to the stored template, then the right decipher key will be regenerated, immediately used to decipher the message (this is the service) and destroyed (the key never leaves the secure area). If the extracted minutiae are not the genuine minutiae, a key is still generated but not the correct key. The message is then incorrectly deciphered, giving a meaningless result. At the end, there is no information revealed, which is a very good property of a secure system, and it is not possible to apply a scheme such as hill-climbing, based on access of the matching score.

It is possible to reach these objectives, but it is hard to achieve and to prove, especially for fingerprints.

Cryptography is Accurate; Biometrics is Fuzzy

In cancelable biometrics, a function similar to a hash function had to be applied, but the data variability was a problem. Intricated biometric involves re-generating a cryptographic key and the same problem. Every bit must be correct; no error is allowed. With biometrics, there is always some uncertainty. Each time a fingerprint is scanned or applied to a sensor, it may not be exactly the same area. The person may have a new cut or scar; the finger could be dirty, wet, or dry.

A partial solution would be to extract a stable sequence from a fingerprint image, always the same, and then combine it with a cryptographic key. This is like extracting a stable signal from a noisy, fuzzy environment. Some research proposed the use of error-correcting code, with the so-called "fuzzy extractor" [8, 9] that can be applied to different biometric modalities, and then specifically over fingerprint databases [10, 11].

The "► fuzzy vault" was proposed in 2002 by Juels et al. [12]. The proposal involved secret being merged with biometric data such as minutiae that does not need to be in a specific order. ► Chaff points [13] are added to hide the genuine minutiae. The experiment was later enhanced using lattice [14], tested on the FVC-2002 database and enhanced with helper data by Uludag et al. [15, 16].

Soutar et al. [2] proposed in 1999 using filters to extract stable characteristics of the fingerprint and then merged them with a secret.

One example scheme is:

- Enroll
  - A set of M minutiae is extracted from a fingerprint.
  - A secret key is divided into M pieces of data; each piece is linked to one minutiae.
  - Random chaff points are added, corresponding to non-existing minutiae and wrong pieces of secret key.
- Recognition
  - A live set of N minutiae are extracted from a live fingerprint.
  - The matching minutiae enable extraction of the correct piece of the secret key.

As the live minutiae may not be exactly the same, it is important to introduce some kind of redundancy for the secret key. A subset of the M enrolled minutiae is needed to perform a match. Lagrange interpolation has been proposed to recover the full secret key, with the advantage of not depending on the order of the points or minutiae.

But some problems arise:

- *Brute force attack:* It is important to add enough chaff points to hide the genuine points and to create too many possible combinations for a brute force attack to succeed.
  - Generating chaff points is not a simple operation, because care must be taken to avoid flaws. The chaff points must be indistinguishable from genuine points. It is a similar problem to random number generators, where it is difficult to prove that they are really random. Always using the same chaff points would make it too easy to find them.
  - Matching minutiae for key extraction will likely require more computation.
  - Chaff points may lead to wrong alignments, especially with poor fingerprints, making minutiae matching less robust.
  - Intricated fingerprint template requires more memory space than a simple template (but another key and program would be needed for protection).

Conclusion

Fingerprint hashing (intricated biometrics) seems to be the ultimate protection scheme. This is not a proven technology yet, but achieving the objectives would lead

to a better protection of privacy without worrying about databases.

Related Entries

- ▶ Encryption, Biometric
- ▶ Fake Finger Detection
- ▶ Fingerprint Features
- ▶ Fingerprint Matching, Automatic
- ▶ Fingerprint Templates

References

1. Cavoukian, A., Stoianov, A.: Biometric encryption: A positive-sum technology that achieves strong authentication, security and privacy. White paper, Information and privacy commissioner of Ontario, March (2007)
2. Soutar, C., Roberge, D., Stoianov, A., Gilroy, R., Vijaya Kumar, B.V.K.: Biometric Encryption, chap. 22, McGraw-Hill (1999)
3. Ratha, N.K., Connell, J.H., Bolle, R.M.: Enhancing security and privacy in biometrics-based authentication systems. IBM Syst. J. **40**(3), 614–634 (2001)
4. Davida, G.I., Frankel, Y., Matt, B.J., Peralta, R.: On the relation of error correction and cryptography to an off-line biometric based identification scheme. In: Proceedings of the Workshop on Coding and Cryptography, Paris, France, pp. 129–138 (1999)
5. Linnartz, J.P., Tuyls, P.: New shielding functions to enhance privacy and prevent misuse of biometric templates. In: Proceedings of the Fourth International Conference on Audio and Video based Biometric Person Authentication, Guildford, UK, pp. 393–402 (2003)
6. Lumini, A., Nanni, L.: An improved biohashing for human authentication. Pattern Recognit. **40**, 1057–4065 (2007)
7. Boult, T.E., Scheirer, W.J., Woodworth, R.: Revocable Fingerprint Biotokens: Accuracy and Security Analysis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07), Minneapolis, USA, pp. 1–8, 17–22 June (2007)
8. Dodis, Y., Reyzin, L., Smith, A.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In: Proceedings of the Eurocrypt 2004, pp. 523–540 (2004)
9. Burnett, A., Byrne, F., Dowling, T., Dury, A.: A biometric identity based signature scheme. In: Proceedings of the Applied Cryptography and Network Security Conference, New York, USA (2005)
10. Costanzo, C.R.: Biometric cryptography: Key generation using feature and parametric aggregation. Online techreport, School of Engineering and Applied Sciences, Department of Computer Science, The George Washington University, October (2004)
11. Al-Tarawneh, M.S., Khor, L.C., Woo, W.L., Dlay, S.S.: Crypto key generation using contour graph algorithm. In: Proceedings of the 24th IASTED International Multi-Conference Signal Processing, Pattern Recognition and Applications, Insbruck, Austria, February (2005)
12. Juels, A., Sudan, M.: A fuzzy vault scheme. In: Lapidoth, A., Teletar, E. (eds.) Proceedings of the IEEE International Symposium on Information Theory, p. 408. IEEE Press (2002)
13. Chang, E.-C., Li, Q.: Hiding secret points amidst Chaff. In: Proceedings of the Eurocrypt, Saint Petersburg, Russia (2006)
14. Zheng, G., Li, W., Zhan, C.: Cryptographic key generation from biometric data using lattice mapping. In: Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), Washington, DC, USA, pp. 513–516. IEEE Computer Society (2006)
15. Uludag, U., Jain, A.K.: Fuzzy fingerprint vault. In: Proceedings on Workshop: Biometrics: Challenges Arising from Theory to Practice, August 2004, pp. 13–16 (2004)
16. Uludag, U., Jain, A.: Securing fingerprint template: Fuzzy vault with helper data. In: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop, June 2006, pp. 163–170 (2006)

First Level Detail

This reflects the general flow of the papillary ridges which may form certain patterns such as arches, loops, whorls, and deltas.

▶ Fingerprint Matching, Manual

Fisher Criterion

Fisher criterion is a discriminant criterion function that was first presented by Fisher in 1936. It is defined by the ratio of the between-class scatter to the within-class scatter. By maximizing this criterion, one can obtain an optimal discriminant projection axis. After the sample being projected on to this projection axis, the within-class scatter is minimized and the between-class scatter is maximized.

▶ Non-linear Techniques for Dimension Reduction

Fixed Pattern Noise

It is characterized by the same pattern of "hot" pixels occurring with images taken under the same conditions of temperature and exposure.

▶ Face Device

Focal Distance

The distance that is required between the iris acquisition device and the iris, for the system to be able to acquire and accurately recognize.

► Iris Acquisition Device

Focal Length

With respect to a lens or mirror, the distance from the lens or mirror at which a parallel beam of light rays will be focused to the smallest size possible for the lens or mirror. The focal length of a simple converging (convex) lens can be measured by focusing the rays from the sun to the smallest point possible and measuring the distance from the image to the lens.

► Face Device
► Iris Device

Footprint Comparison

► Forensic Barefoot Comparisons

Footstep Identification

► Footstep Recognition

Footstep Recognition

Ruben Vera Rodriguez[1], Nicholas W. D. Evans[1,2], John S. D. Mason[1]
[1]Swansea University, Singleton Park, Swansea, SA2 8PP, UK
[2]Institut Eurécom, 2229 rote des Crêtes, 06560 Sophia-Antipolis, France

Synonyms

Footstep identification; Footstep verification

Definition

Footstep recognition is a relatively new biometric and is based on the study of footstep signals captured from persons walking over an instrumented sensing area. Since the biometric information is embedded in a time varying signal, thereby implying some form of action (in this case those of walking or running for example), footsteps can be included in the group of behavioral biometrics.

Introduction

Footstep recognition was first suggested as a biometric in 1977 by Pedotti [1], but it was not until 1997 when Addlesee et al. [2] reported the first experiments. Since then the subject has received relatively little attention in the literature and so it is perhaps of little surprise that reported performances fall short of those achievable with other, more popular, and researched biometrics. However, recent work has demonstrated the real potential of the footstep biometric which is certainly not without its appeal.

One significant benefit of footsteps over other, better known biometrics is that footstep signals can be collected covertly with minimal client cooperation. Other benefits lie in the robustness to environmental noise (a limiting aspect of speaker recognition) or lighting variability (as in the case of face recognition). There is, however, a number of new challenges to be addressed. Footsteps can exhibit a high degree of intra-class variability, i.e., different footwear, persons carrying heavy baggage and different walking speeds, all

extraneous factors which make footstep recognition an extremely challenging task.

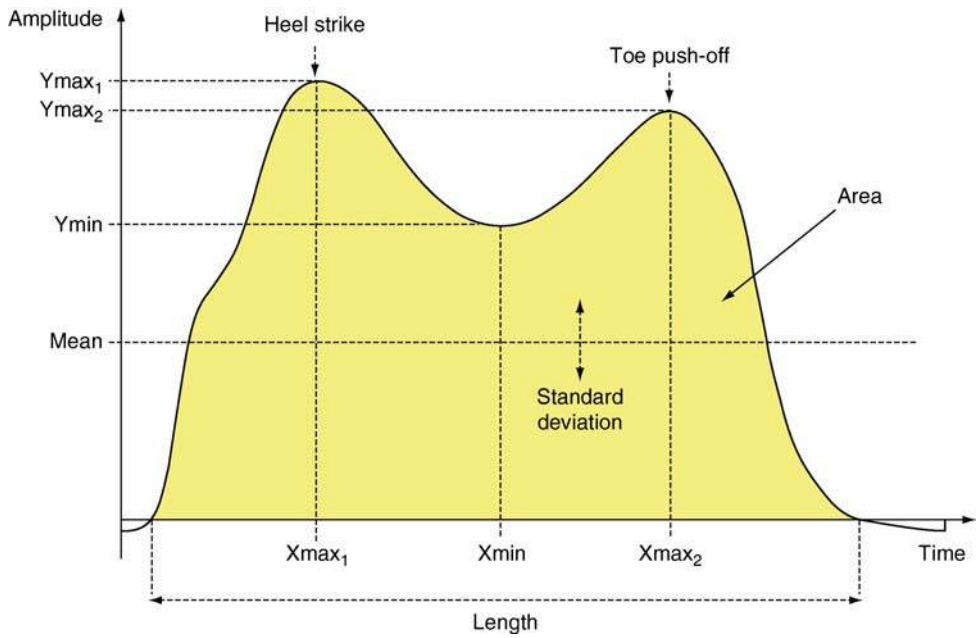
In addressing these difficulties among others, researchers have investigated footstep signals using different sensor approaches. Systems reported in the literature include the extraction of footstep positions using video cameras, acoustic-based approaches which capture the sound of footsteps [3] and, by far the most common, under-floor contact or tactile-based sensors. These approaches range from simple ON/OFF sensors that indicate the position of the footstep [4–7] to more sophisticated sensors that capture transient pressure [1, 2, 8–13]. Pressure sensors generally measure the ground reaction force (▶ GRF). An example GRF profile for a single footstep signal captured from the sensor approach reported in [13] is shown in Fig. 1. Generally there are two peaks to the GRF profile, the first peak is attributable to the heel strike and the second to the toe push-off as the body is propelled forward. Figure 1 also illustrates some of the most common geometric features (maximum, minimum and mean values) as used in the works of [9, 12, 14] for subsequent classification.

Reported performances vary widely. The most statistically meaningful results obtained for footstep recognition with an identification protocol relate to a database comprised of 1,680 footsteps from 15 persons [9]. Here an accuracy of 93% was reported. For the case of verification as a protocol, best results relate to a database comprised of 3,147 footsteps from 41 persons [15]. Equal error rates (EERs) of 9.5 and 13.5% are reported for development and evaluation sets respectively. Results to date are promising and show that the use of footsteps as a biometric warrants further investigation.

The following sections present an overview of different applications of footstep signals and a review of published literature which has investigated the use of footsteps more specifically for biometrics.

Applications

It is possible to classify different biometric techniques according to the original application of the biometric signal. In the case of the fingerprint and hand geometry biometrics, signals are captured with the sole application of biometrics; whereas for speech, for example, the main application is communication, and biometrics can be considered a secondary application. Other biometrics such as the footsteps are in the middle of this range. A footstep is an action that can be captured for several applications. Potential uses of footstep signals in the



Footstep Recognition. Figure 1 Example of a GRF profile against time for a single footstep. The first peak corresponds to the heel strike and the second corresponds to the toe push-off.

literature include medicine, surveillance, smart homes, multimedia, and biometrics, none of them dominating and therefore this overview presents the entire spectrum.

In the field of medicine, footstep signals have been used to analyse different gait deficiencies by comparing normal and pathological patterns of footstep pressure signals. Following early work on biomechanics, in 1977 Pedotti [1] studied the three orthogonal components of the GRF signal using a square force plate with four piezoelectric transducers placed in the corners, similar to other systems used later for biometrics [2, 9, 10, 12]. He studied visually around 4,500 footsteps from 65 normal and 165 pathological subjects and observed stride symmetry between the left and the right feet for normal subjects but not for pathological subjects; furthermore, Pedotti noted low intra-person variability, leading to one of the first suggestions to the use of footsteps as a biometric. Commercial products today provide high resolution pressure image sequences from thin sensor mats created by printing processes. These systems are used in medicine to study for example the plantar pressure profiles, identify asymmetries between left and right feet, review dynamic weight transfer and local pressure concentrations, or identify areas of potential ulceration amongst others.

More focused on the detection of footsteps for surveillance applications, footstep signals have been used to detect human presence in a determined area. The work described in [3] reports some experiments carried out with a database comprised of five people walking ten times toward a microphone. The aim of the research was not only on footstep detection but person identification using mel-cepstrum analysis. Other work reported in [16] used piezoelectric accelerometers to detect impulses induced by walking. Footsteps were identified from three or more impulses where the sensor was excited at its resonant frequency, having satisfactory results in most occasions.

One particularly appealing application of footstep signals is found in the field of smart homes. In 2000 Mori et al. [17] developed a robotic room where multiple sensors were distributed in several locations. Footstep signals were collected from a distribution of force sensing resistors (FSRs) to specify human position in the room. A total number of 252 FSRs were installed in a 200 mm × 200 mm lattice shape. More recent work on the same floor [4] (2002) increased the spatial resolution of the sensors to a 64 × 64 switch sensor array in a 500 mm^2 space. With this higher resolution, experiments determined the positions of a human and a four-wheeled cart and distinguished between them. In 2004 Murakita et al. [5] reported a system for tracking individuals over an area of 37 m^2 employing basic block sensors of 18 cm^2. The system was capable of tracking two different people when separated by more than 1.4 m but failed to track people in a crowded area due to the low spatial resolution and a low capture rate of 5 Hz. Making use of the hardware developed for the Active Floor [2], in 2001 Headon and Curwen [18] used the vertical component of the GRF and a hidden Markov model (HMM) classifier to recognise different movements including stepping, jumping, or sitting down. Applications of such a system exist in safety (i.e., fall detection for the elderly) and entertainment (i.e., video games). More recently, in 2008 Liau et al. [19] developed a system which used load cells over an area of 4 m × 4 m to track people and addressed the cross-walking problem where the paths of two or more people intersect.

Footstep signals have also been used for multimedia applications. In 1997 Paradiso et al. [20] developed a system which he called The magic carpet to be used in an audio installation where users created and modified complex musical sounds and sequences as they wandered about the carpet. The sensor floor comprised a 16 × 32 grid of piezoelectric wires in an area of 1.8 m × 3 m carpet. Later in the same year, the same laboratory developed a system installing PVDF (polyvinylidene fluoride) and FSR sensors into a dancing shoe [21]. The goal was to capture many degrees of expression and use them to drive music synthesizers and computer graphics in a real-time performance. More recently, in 2005 Srinivasan et al. [8] developed a portable pressure sensing floor constructed of modular high resolution pressure sensing mats. A sensor mat comprised 2,016 sensors made from a pressure sensitive polymer and covered an area of 62 cm × 53 cm, sampling each sensor at a frequency of 30 Hz. Initial applications of the system were to study interactive dance movement and video game controlling.

Review of Footsteps as a Biometric

Review of footsteps as a biometric is now an addressed work in the open literature which considers the use of footsteps specifically as a biometric. One of the first investigations into footstep recognition was reported

by UK researchers in 1997 [2]. They reported experiments on a database of 300 footsteps signals that were captured from 15 walkers in one session. The system was comprised of four load cells measuring the vertical component of the GRF and placed on the corners of a tile working at a sampling frequency of 250 Hz. They divided the database into train and test and an identification accuracy of 91% was achieved with an HMM classifier and samples from the GRF of a single footstep signal as features.

In 2000, and using a similar sensor approach, a group in the USA reported results on a database of 1,680 footstep signals collected from 15 persons using a frequency sampling of 150 Hz [9]. Signals were collected from both left and right feet and different footwear having 20 footsteps per condition using half of them for training and half for testing. Ten geometric features were extracted from the GRF of a single footstep signal including the mean value, the standard deviation, maxima, and minima, etc. They considered each combination of user, foot, and shoe type as a cluster. Then a nearest neighbour classifier was used to measure the Euclidean distance of a footstep from the test set to each cluster. An identification accuracy of 93% was reported regardless of whether the correct shoe or foot was given. In 88% of the cases, a user's footstep was more similar to other footsteps for that same user than for another user, concluding from these results that footwear does not greatly affect the ability of their approach to identify the user by his footsteps.

While focused toward the study of gait, a group from Switzerland [10] developed in 2002 a system fusing data acquired from 3 tiles of 4 piezo force sensors each and video cameras. A database of 480 footsteps was collected from 16 persons walking barefoot using a sampling frequency of 300 Hz. The database was further divided into train and test. They studied different feature extraction techniques as geometric features from GRF [9] and phase plane (as area within the curve, position of the loop, maxima, minima, etc.). The best verification performance was achieved using the power spectral density (PSD) of the derivative GRF of footsteps signals in the band of 0–20 Hz with generalized principal component analysis (GPCA), obtaining a verification EER of 9.5% with an Euclidean distance classifier.

A Korean group reported a system in 2003 [6] that used 144 simple ON/OFF switch sensors in a total area of 1 m × 3 m. Stride data (connected footsteps) was collected from ten persons who each one walked 50 times across the ubiFloor resulting in a database of 500 walking samples. Then the database was divided into training, validation, and testing data randomly. The position of several connected footsteps was used as users walking features instead of the pressure of one footstep, as proposed in [2, 9]. An accuracy of 92% was reported with a multi-layer perceptron (MLP) neural network used as an experimental identification method.

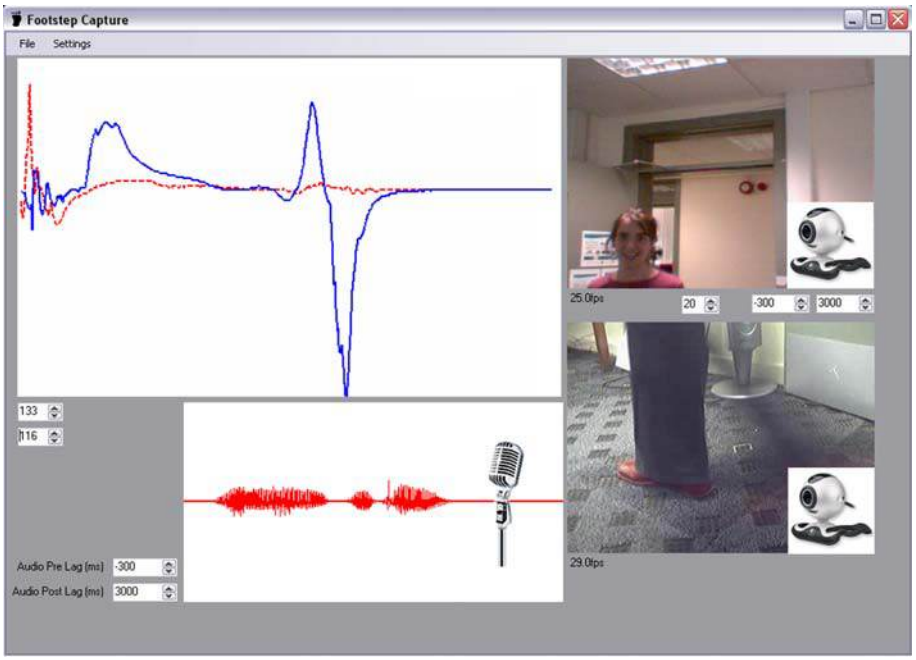
In 2004 a group from Finland investigated footstep recognition using electro mechanical film (EMFi). Long strips of the sensor material were laid over an area covering 100 m^2. A database of around 440 footstep signals (of both feet) was collected from 11 persons at a frequency rate of 100 Hz. In their publication [11] they reported experiments with a two level learning vector quantisation (LVQ) based classifier and considered three consecutive footsteps of a person to carry out a single test. On the first level each of the three single footstep signals was classified independently, and on the second level the decisions of the three consecutive footsteps were taken into account having a final acceptance if a majority of the footsteps were classified to the same class. The recognition rate reported was 89% of accuracy with an 18% of rejection rate. In the same year they reported different experiments [14] based on the same database. Geometric features were extracted from the GRF profiles as in [9] and first FFT coefficients. Using a distinction-sensitive LVQ (DSLVQ) classifier for a single footstep, an identification accuracy of 70% was achieved. Later in 2005, they presented experiments in [22] combining different feature sets using a two level classifier. On the first level three different feature sets were extracted from a single footstep as geometric features from the GRF as in [14], FFT of GRF with PCA, and FFT of the derivative GRF with PCA. Then, a product rule was used to combine the three results obtained. On the second level different footsteps from the same person were combined using an average strategy. These experiments were done for two classifiers: LVQ and a MLP neural network. Results were better for MLP classifier in all cases, having a recognition rate of 79% for the case of a single footstep and a 92% for three consecutive footsteps.

In 2005 a group from Southampton (UK) [7] reported trials with a system comprising 1,536 sensors arranged in a 3 m × 0.5 m rectangular strip with an individual sensor area of 3 cm^2. A database of 180

signals was collected from 15 people without wearing footwear at a frequency of 22 Hz. Each person walked over the mat 12 times and in each case two complete gait cycles (4 foot falls) were captured. Three features were extracted: stride length, stride cadence, and heel-to-toe ratio. An identification accuracy of 80% was reported using a nearest neighbor classifier to measure the Euclidean distance between each feature vector and the mean feature vector of the experimental population, i.e., the whole database. This work along with the early work of [6], differs from other published material in using binary signals rather than sampled waveforms and capture stride information from a short series of footfalls. Stride characteristics are also considered by [11, 22] as stated above.

In 2006 another group from Southampton [12] investigated a system similar to the work in [2, 9]. A database of 400 signals was collected from 11 people. Using geometric features extracted from GRF profiles as in [9] an identification accuracy of 94% was achieved using a nearest neighbor classifier in the same way as in [7].

More recently, in 2007, a research group from Swansea (UK) presented in [13, 15] experiments obtained with a database comprised of 3,174 footsteps from 41 different persons in different sessions and shoes from two piezoelectric transducers sampled at a frequency of 1,024 Hz. The database was further divided into independent development and evaluation datasets adopting a standard best practice evaluation strategy, and therefore, presenting more statistically meaningful results and potentially more reliable predictions of performance. The database is freely available to the research community [23]. Due to the amount of data collected, a semi-automatic footstep capture system was developed to facilitate automatic labeling and rapid manual validation. Figure 2 shows a screenshot of the footstep capture system user interface. A microphone captured a spoken ID used for automatic speaker recognition to label the data (bottom part of Fig. 2); and two video cameras, one recording the face and the other the foot (top and bottom right part of Fig. 2 respectively), were used for manual data validation; the sensor responses are illustrated in the top left part of Fig. 2 as a function of time (horizontal axis). For feature extraction, two approaches were followed, namely geometric and holistic. The geometric approach was based on the extraction of main characteristic points of the footstep profile: the area, mean, length, maxima/minima, etc. The holistic approach was based on both sensor outputs and the GRF profile after PCA to reduce dimensionality of the data. In [13] two different



Footstep Recognition. Figure 2 Screenshot of the footstep capture system software developed in [13, 15].

classifiers, a nearest neighbor and SVM were also compared and findings were as expected that SVM outperforms the NN, and surprisingly holistic features outperforms the geometric features. Results of 9.5 and 11.5% EER were obtained for development and evaluation sets respectively for holistic features with an SVM classifier. Following best-practice, a formal assessment protocol was defined for the footstep recognition evaluation presented in [15]. The protocol reflects that utilized by the international NIST speaker recognition evaluations. Also, an optimization of the two feature approaches was carried out obtaining results of 9.5% EER for the development set and 13.5% EER for the evaluation set using optimized holistic features with an SVM classifier. EER given of 13.5% corresponds to 1,697 errors of each class (false acceptance and false rejection) from a total number of 25,143 tests. Such simple analysis allowing comparison across systems comes from adopting the task with verification. Work is ongoing with a multi-sensor stride capture system with the primary goal of improving confidence in the assessment of footsteps as a biometric.

Table 1 presents a comparison of this related work. The second column shows that relatively small database sizes is a common characteristic of the earlier work certainly judged in relation to other biometric evaluations where persons are normally counted in hundreds or thousands and the number of tests perhaps in many thousands. A maximum number of 16 persons and 1,680 footstep examples were gathered in all cases except in [13, 15] which reports results on 3,147 footsteps and 41 persons. In each case, except for [7, 12], the databases are divided into training and testing sets, but none use independent development and evaluation sets, with exception of [13, 15], a limitation which makes performance predictions both difficult and unreliable. Identification, rather than verification, was the task considered in all but three of the cases, the exceptions being [10, 13, 15]. Identification has the benefit of utilizing the available data to a maximum but suffers from well known scalability problems in terms of the number of classes in the set. Also, it is interesting to point out that some systems present classification results for stride data (consecutive

Footstep Recognition. **Table 1** A comparison of different approaches to footstep recognition 1997–2007

Group, year	Database (total steps/ persons)	Technology	Features	Classifier	Results
The ORL Active Floor (UK) 1997 [2]	300 steps, 15 persons	Load cells	Sub sampled GRF	HMM	ID rate: 91%
The Smart Floor (USA) 2000 [9]	1,680 steps, 15 persons	Load cells	Geometric from GRF	NN	ID rate: 93%
ETH Zurich (Switzerland) 2002 [10]	480 steps, 16 persons	Piezo force sensors	Power spectral density	Euclidean density	Verif EER: 9.5%
Ubifloor (Korea) 2003 [6]	500 steps, 10 persons	Switch sensors	Position of several steps	MLP neural net.	ID rate: 92%
EMFi Floor (Finland) 2004 [11, 14, 22]	440 steps, 11 persons	Electro mechanical film	Geometric from GRF, and FFT	MLP neural net. and LVQ	Best ID rate [22] of 92% using three footsteps as test
Southampton University (UK) 2005 [7]	180 steps, 15 persons	Resistive (switch) sensors	Stride length, cadence and heel-to-toe ratio	Euclidean distance	ID rate: 80%
Southampton University (UK) 2006 [12]	400 steps, 11 persons	Load cells	Geometric from GRF	NN	ID rate: 94%
Swansea University (UK) 2007 [13, 15]	3,174 steps, 41 persons	Piezoelectric sensors	Geometric and holistic	SVM	[15] Verif EER: 9.5% for Devel, 13.5% for Eval

footsteps) [6, 7, 11, 14, 22] while the rest only for a single footstep [2, 9, 10, 12, 13, 15]. In [22] an identification accuracy of 79% using a single footstep as a test was improved to 92% when three consecutive footsteps were used. This equates to a relative improvement of 16%.

Summary

Footstep recognition is a relatively new biometric relative to other biometrics in terms of the research reported in the literature. As reviewed, footstep signals have been used for different applications, thus different capture systems have been developed. In the field of biometrics the same trend is observed; researchers have developed systems with different sensors, extracting different features, and with different assessment protocols. Recently, in 2007, the world's first freely available footstep database was released to the research community [23]. Of particular importance to this development is, not only the size of the database both in terms of the number of footsteps and clients, but the standard, best practice evaluation protocols that accompany the database. For the first time researchers will be able to develop and assess new approaches on a common and meaningfully sized database. As has happened for many other biometric modalities, it is hoped that this will stimulate new interest in the footstep biometric, lower the cost of entry and provide a solid foundation for future research.

Given its current state of development the future of footstep recognition research is difficult to predict. Some obvious avenues include new features and novel normalization approaches to reduce the effects of extraneous factors. Other possibilities include further investigation into connected footsteps, i.e., stride information, information that isn't captured by single footstep systems. This research would explore the middle ground between footsteps and gait. Gait is another biometric that finds applications in different areas such as in medicine, the sports industry, and biometrics. In the biometrics context, gait aims to recognise persons from a distance using walking characteristics extracted from video recordings. In contrast, footsteps are a more controlled biometric due to the fixed, constrained sensing area. It would thus seem natural for future research to investigate the fusion of the two biometrics.

Related Entries

▶ Gait Recognition

References

1. Pedotti, A.: Simple equipment used in clinical practice for evaluation of locomotion. IEEE Trans. Biomed. Eng. **BME-24**(5), 456–461 (1977)
2. Addlesee, M.D., Jones, A., Livesey, F., Samaria, F.: The ORL active floor. IEEE Pers. Commun. **4**(5), 35–41 (1997)
3. Shoji, Y., Takasuka, T., Yasukawa, H.: Personal identification using footstep detection. In: Proceedings of 2004 International Symposium on Intelligent Signal Processing and Communication Systems, pp. 43–47 (2004)
4. Morishita, H., Fukui, R., Sato, T.: High resolution pressure sensor distributed floor for future human–robot symbiosis environments. In: Proceedings of 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 2, pp. 1246–1251 (2002)
5. Murakita, T., Ikeda, T., Ishiguro, H.: Human tracking using floor sensors based on the Markov chain Monte Carlo method. In: Proceedings of the 17th International Conference on Pattern Recognition (ICPR), vol. 4, pp. 917–920 (2004)
6. Yun, J.S., Lee, S.H., Woo, W.T., Ryu, J.H.: The user identification system using walking pattern over the ubiFloor. In: Proceedings of International Conference on Control, Automation, and Systems, pp. 1046–1050 (2003)
7. Middleton, L., Buss, A.A., Bazin, A.I., Nixon, M.S.: A floor sensor system for gait recognition. In: Proceedings of Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05), pp. 171–176 (2005)
8. Srinivasan, P., Birchefield, D., Qian, G., Kidane, A.: A pressure sensing floor for interactive media applications. In: Proceedings of the 2005 ACM SIGCHI International Conference, vol. 265, pp. 278–281 (2005)
9. Orr, R.J., Abowd, G.D.: The smart floor: a mechanism for natural user identification and tracking. In: Proceedings of Conference on Human Factors in Computing Systems, pp. 275–276 (2000)
10. Cattin, C.: Biometric authentication system using human Gait. Swiss Federal Institute of Technology, Zurich. PhD Thesis (2002)
11. Suutala, J., Pirttikangas, S., Riekki, J., Roning, J.: Reject-optional LVQ-based two-level classifier to improve reliability in footstep identification. Lecture Notes Comput. Sci. Springer, Berlin **3001**, 182–187 (2004)
12. Gao, Y., Brennan, M.J., Mace, B.R., Muggleton, J.M.: Person recognition by measuring the ground reaction force due to a footstep. In: Proceedings of Ninth International Conference on Recent Advances in Structural Dynamics (2006)
13. Vera-Rodriguez, R., Evans, N.W.D., Lewis, R.P., Fauve, B., Mason, J.S.D.: An experimental study on the feasibility of footsteps as a biometric. In: Proceedings of 15th European Signal Processing Conference (EUSIPCO'07), pp. 748–752. Poznan, Poland (2007)

14. Suutala, J., Roning, J.: Towards the adaptive identification of walkers: automated feature selection of footsteps using distinction-sensitive LVQ. In: Proceedings of International Workshop on Processing Sensory Information for Proactive Systems, pp. 61–67 (2004)
15. Vera-Rodriguez, R., Lewis, R.P., Evans, N.W.D., Mason, J.S.D.: Optimisation of geometric and holistic feature extraction approaches for a footstep biometric verification system. In: Proceedings International Summer School for Advanced Studies on Biometrics for Secure Authentication. Alghero, Italy (2007)
16. Mazarakis, G.P., Avaritsiotis, J.N.: A prototype sensor node for footstep detection. In: Proceedings of the Second European Workshop on Wireless Sensor Networks, pp. 415–418 (2005)
17. Mori, T., Sato, T., Asaki, K., Yoshimoto, Y., Kishimoto, Y.: One-room-type sensing system for recognition and accumulation of human behavior. In: Proceedings of 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 1, pp. 344–350 (2000)
18. Headon, R., Curwen, R.: Recognizing movements from the ground reaction force. In: Proceedings of the 2001 Workshop on Perceptive User Interfaces, vol. 15, pp. 1–8. Orlando, USA (2001)
19. Liau, W.H., Wu, C.L., Fu, L.C.: Inhabitants tracking system in a cluttered home environment via floor load sensors. IEEE Trans. Autom. Sci. Eng. **5**(1), 10–20 (2008)
20. Paradiso, J., Abler, C., Hsiao, K., Reynolds, M.: The magic carpet: physical sensing for immersive environments. In: Proceedings of CHI'97, pp. 277–278. Atlanta, USA (1997)
21. Paradiso, J., Hu, E.: Expressive footwear for computer-augmented dance performance. In: Proceedings of the First international Symposium on Wearable Computers. IEEE Computers Society Press, pp. 165–166. Cambridge, USA (1997)
22. Suutala, J., Roning, J.: Combining classifiers with different footstep feature sets and multiple samples for person identification. In: Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 5, pp. 357–360 (2005)
23. S.U.: Footstep recognition at Swansea University. Available at http://eeswan.swan.ac.uk

Footstep Verification

► Footstep Recognition

Footwear Marks

Footwear marks is an umbrella term describing the various types of marks that an item of footwear can produce through its use. The terms outsole print, imprint and impression, or footwear print, imprint and impression are collectively called footwear marks.

► Footwear Recognition

Footwear Recognition

Maria Pavlou, Nigel M. Allinson
University of Sheffield, Mappin Street, Sheffield, UK

Synonyms

Outsole pattern matching; Shoeprint matching

Definition

Footwear recognition is the process of acquiring, identifying, and verifying the marks of the outsole (underside) patterns of a shoe. These marks arise as a result of the normal use of footwear in many conditions and environments. Footwear recognition can be used by the police and other law enforcement agencies in the identification of crime suspects.

Introduction

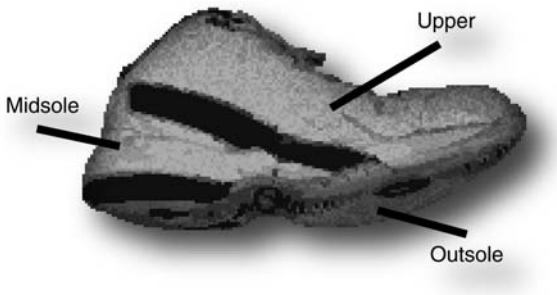
Although footwear recognition in a strict sense is not a biometric, it does provide a very useful source of intelligence and potential evidence in the application of forensics for policing and security. As shoes are fairly personal items of apparel with usually an extended period of ownership by their wearer, they could be termed a "near-biometric." Similar to latent fingerprints, ► footwear marks are very frequently left behind on surfaces at crime scenes [1]; and they can be more commonly recovered than fingerprints for some crime categories. A number of methods are then used to develop and collect these ► scene marks to provide useful evidential clues by linking patterns of movement of suspect individuals (at crime scenes), and can even provide strong courtroom evidence by matching a mark to an individual shoe. This useful resource has gained recent interest internationally, even resulting in

legislative changes in the UK [2] where collected footwear evidence is treated in the same way as fingerprint and DNA evidence. Namely, they have to be provided at time of arrest, and can be held and searched on local/national computer systems.

Etiology, Detection and Recovery

The typical shoe comprises of three parts (see Fig. 1) – upper, midsole, and outsole. The footwear upper is generally constructed from a variety of hard wearing fabrics or leather and is fashioned and colored in a multitude of ways. The upper holds the foot firmly in place and provides suitable support. The midsole holds the inner sole and is also used to fasten the uppers to the outsole. The outsole is the underside of the footwear, made of a durable leather, rubber or polyplastic, which provides traction and cushioning for the wearer. Manufacturers have made great efforts in the design of the outsole for the benefit of the wearer in varying activities by incorporating functional and decorative ► tread patterns. More commonly worn leisure footwear or sneakers typically have intricate tread designs based on the shoe model theme or manufacturer logos. What is important here is that the tread pattern is usually very distinctive to any design of shoe model just like the friction skin ridge patterns of fingers are unique to an individual.

Similar to latent fingerprints, it is the contact of the outsole with various surfaces that results in the formation of a footwear mark in a number of ways. This can be from the deposition of dry material such as dust or dirt, or wet materials such as water, blood or mud, onto a surface. The removal of material from a surface may also form a mark, leaving a negative impression for example when stepping into and out of a shallow pool of blood, while an indented impression can be formed in a soft substrate such as snow or clay. Accordingly for each type of mark there are numerous methods by which these are detected, recorded, and preserved. Details can be found in [1, 3]. Briefly, these range from using specialized lighting methods, such as oblique and multispectral lighting, and chemical developers to enhance hard to see traces which can then be photographed. Several lifting techniques are also used to capture deposited particle materials onto a fixing substrate such as a ► gelatin pad. When a footwear mark is left in a soft material, such as snow, specialized plaster or molten sulphur can be used to produce a cast of the impression. Finally a print of the outsole can be made directly if available. This is done using dusting techniques, such as fine aluminum powder and then pressing onto a transparent gel sheet. More commonly the outsole can be impregnated with a dye, and printed onto paper, or with an oil-based liquid and printed on special sensitized paper – a method commonly called Printscan (see Fig. 2). This last method is the technique most employed in Police ► custody suites to produce impressions of a suspect's shoes. The resulting impression can then be used for one-to-one comparisons to provide forensic evidence, or can be scanned for computer-based processing. The overriding aim of all these development and recovery techniques is to obtain as true and unaltered a representation of the mark as possible for later processing and examination.



Footwear Recognition. Figure 1 Components of a typical athletic shoe, comprising the upper, midsole, and outsole.

Uniqueness and Application

Intuitively the uppers of an item of footwear are immediately more useful in identifying the make or model of a shoe. This is because of their styling, coloring, and the presence of manufacturer logos, with detailed information being readily obtainable from outlets and manufactures. However the uppers and any associated markings are rarely encountered as forensic evidence. The impressions and marks produced by the outsole are more readily found and can contain sufficient characteristic information to ascertain the manufacturer, model, and potentially the wearer. These characteristics originate in a number of

Footwear Recognition. Figure 2 Making an outsole pattern print on sensitized paper with the Printscan method.

ways starting from the manufacture process. Footwear manufactures use a number of processes for the production of outsoles [1] resulting in a varying degree of ► process artifacts and defects remaining in the final product. These are one of the three useful characteristics of an outsole, which also comprises the outsole tread pattern and the accumulated wear-and-tear artifacts. Of these characteristics only the wear-and-tear artifacts are unique provided they have occurred due to a random process where something is added or taken away from the outsole that either causes or contributes to making the outsole unique. Such artifacts include nicks, cuts, scratches and ► feathering of the rubber material due to the normal usage of the footwear. These can be called "individual characteristics" while the outsole tread pattern and other manufacture defects are termed "class characteristics" which are distinct to a particular model of footwear and the process of its production, such as its outsole mold.

The identification and use of these characteristics have different meaning and implications. In a forensic setting the class characteristics are important as they provide information on the manufacture and model of the footwear worn and also its size. This is useful when restricting a suspect list based on physical build/size, accessibility of rare/expensive items and even geographical distribution of crimes. Once candidates of the same footwear class are available only then can comparisons be made on individual characteristics. Forensic examiners will look for common individual characteristics between items of recovered footwear, their reproduced marks and marks found as evidence which provide conclusive links and can be used as court room evidence. Provided there are sufficient individual characteristics between an outsole and a recovered mark it may be possible to state that the outsole created the outsole mark to court room standards of evidence. Outside the forensic setting, footwear class characteristics can be very useful for screening and intelligence gathering. The footwear of a suspect can be collected and from which its class characteristics are ascertained. Usually suspects will be offenders detained or held on an unrelated offense and may have their footwear proactively compared with evidence collected at an earlier time in relation to other offenses in a local area, such as burglaries. If a link is made between an arrestee and marks found at crime scenes then it can provide law enforcement officers with some important information with which to interrogate while the arrestee is still in custody.

Comparison and Identification Methods

The comparison of class and individual characteristics of an outsole have been largely carried out by experienced forensic professionals, as an intimate

knowledge of footwear marks and their etiology is required. However, some efforts have been made to automate the identification of outsole patterns and their associated shoe model based on class characteristics. While this task is feasible, the task of automatically verifying an outsole-to-mark or a mark-to-mark based on class and individual characteristics is much more difficult. Current nation-wide footwear databases in the UK comprise over 15,000 shoe models and are constantly expanding as manufacturers introduce new styles. Crime scene marks are recovered by diverse techniques and the resulting impressions are often of poor quality, confounded by details of the underlying surface and may only represent a partial impression of the entire outsole. Such factors will make the fully automatic identification or verification of outsoles marks a very difficult, if not impossible, option.

Automatic matching of footwear patterns has not been reported much in the literature. Early works [4] have employed semi-automatic methods of ▶ manually annotated footwear print descriptions using a codebook of shape and pattern primitives, for example, wavy lines, geometric shapes, and logos. Searching for an example impression then requires its encoding in a similar manner to that used for the reference database. This process is laborious and can be the source of poor performance as similar patterns may be inconsistently encoded by different users. It is still, however, predominantly used by Police Forces across the UK and elsewhere. Two major factors are attributed to this; firstly, the ease of understanding the coding methodology and its primitives (i.e., their visual intuitiveness), and secondly, a lack of proven and accepted automated systems based on rigorous standards and robust performance.

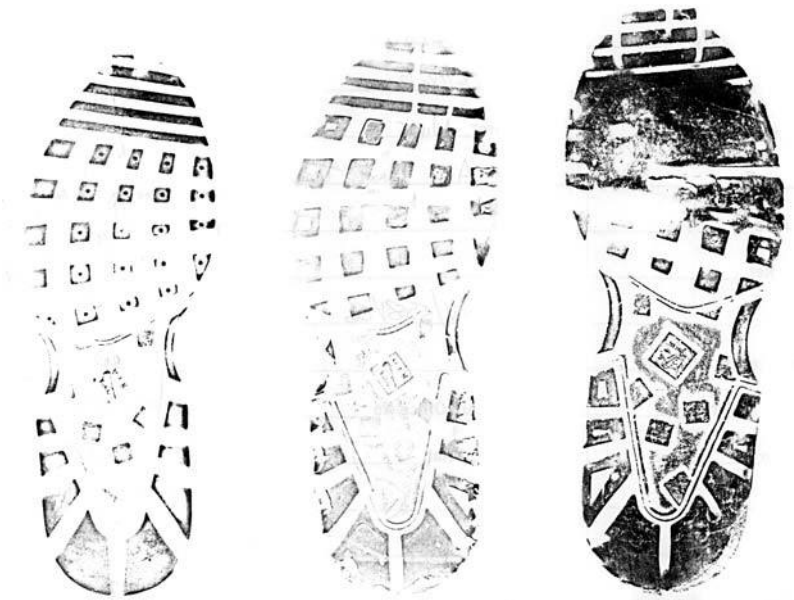
One early work [5] employed an intuitive coding scheme based on shapes automatically generated from footwear images using various ▶ image morphology operators. The spatial positioning and frequencies of these shapes were used for classification with a neural network. Unfortunately, the authors did not report any performance statistics for their system. Later works do not follow this approach but instead use template matching approaches based on the pattern representation in a suitable transform space. Fractal representations were used in [6, 7] with a mean square noise error method for classification. They report good results; however the dataset used was small and contained no spatial or rotational variations. In [8], Fourier Transforms (FT) are used for the classification of full and partial marks of varying quality. The FT provides a degree of invariance to translations and encodes spatial frequency information. By incorporating duplicate rotated templates a degree of rotation invariance was also possible. Their approach was weak on first rank precision, and this may have been due to the large variation in print quality. Also, the footwear prints were processed globally and hence noise in the images could have hindered the quality of useful encoded local information evident in the print. Despite these failings this approach is promising and shows the importance of encoding local information.

In [9] local image features (LIF) were used to make one-to-one comparisons of patterns in a database of footwear marks which had been subjected to added noise and transformed in various ways. This work compared a number of the aforementioned approaches, for which LIF performed very well, indicating that local image information is crucial to good matching performance especially when considering partial marks. However as test samples were generated from the training set it is not clear how this template matching approach would perform on untrained test samples or in a one-to-many and pattern search scenario.

Moving away from a template-based approach, work in [10] employed a histogram approach by quantizing edge directions into 5° intervals. In order to cope with rotational variations of the image, and hence translation of the histogram bins, an FT was also applied. Their approach is useful as now the pattern information is described in a compact way based on the content of edges, however, due to histogram normalization and the use of the FT, this approach is not effective with partial marks.

The flexibility of histogram-based encoding is effective in many image retrieval and indexing tasks as has been demonstrated in [11]. It is interesting that the manual coding schemes still used for footwear impressions bear a strong resemblance to basic histogram coding methods. The work proposed by [12] proceeds by encoding footwear patterns into a compact vector space model using a feature-rich codebook of local feature descriptions. The codebook is a quantized set of features derived by applying such feature detectors

as maximally stable extremal regions (MSERs), followed by the use of robust feature descriptors. A query shoe impression can be coded using this codebook and compared against others. This approach is fully automated and yet still preserves some similarity to manual coding approaches in that the coded features resemble the use of annotated features in the semi-automated systems described previously.



Footwear Recognition. **Figure 3** Outsole marks captured using the Printscan method showing different stages of wear and print quality. From left to right the pattern shown has increasing wear. Note the change in tread pattern appearance.

Footwear Recognition. **Figure 4** Various footwear marks collected from scenes of crime. From left to right, a print in blood on soft tiles, outsole imprint in mud, an electrostatic lift of dirt from carpet, a casting, and a gel lift of a dusty mark.

This approach is able to cope with rotational changes and queries in the form of partial impressions. Though good performance is reported in this study (e.g., for a reference set of 374 different footwear models, a precision at first rank of 87% was obtained), it is difficult to compare the performance of differing approaches as no common datasets have been employed nor is there agreement on testing methodologies.

A number of issues are not addressed by these methods. One is the problem of dealing with large appearance changes in footwear marks as the outsole becomes worn over time and where the mark is strongly degraded or partially missing (see Fig. 3). Some work by Su et al. in [13, 14] looked at methods for assessing and improving image quality of footwear marks collected using the Printscan process. Even if good quality marks are obtainable there is still work to be done on how to compare marks obtained by different methods, for example between a casting and a gel lift (see Fig. 4). Additionally, a lack of standards for the digital capture of marks and the absence of an agreed and openly available dataset are issues that still need to be addressed.

Summary

The use of footwear impressions both recovered from crime scenes and acquired from suspect's shoes have a significant role to play as a "near-biometric" in forensic investigations. It has, despite its long history as an important tool of forensics, remained until recently largely forgotten. This will undoubtedly change in the near future with a much greater development and application of mainstream biometric tools and methodologies.

Related Entries

► Forensic Applications, Overview
► Forensic Barefoot Comparison

References

1. Bodziak, W.J., ed.: Footwear Impression Evidence. CRC Press, Boca Raton (2000)
2. Parliament: Serious Organised Crime and Police Act 2005, Elizabeth II. The Stationery Office (2005)
3. Hilderbrand, D.S., ed.: Footwear, The Missed Evidence. Staggs Publishing (1999)
4. Mikkonen, S., Astikainen, T.: Databased classification system for shoe sole patterns-identification of partial footwear impression found at a scene of crime. J. Forensic Sci. **39**, 1227–1236 (1994)
5. Geradts, Z., Keijzer, J.: The image-database rebezo for shoeprints with developments on automatic classification of shoe outsole designs. Forensic Sci. Int. **82**, 21–31 (1996)
6. Bouridane, A., Alexander, A., Nibouche, M., Crookes, D.: Application of fractals to the detection and classification of shoeprints. In: Proceedings International Conference on Image Processing. Volume 1. pp. 474–477 (2000)
7. Alexander, A., Bouridane, A., Crookes, D.: Automatic classification and recognition of shoeprints. In: Proceedings Seventh International Conference on (Conf Image Processing and Its Applications Publ. No. 465). Volume 2. 638–641 (1999)
8. de Chazal, P., de Chazal, P., Flynn, J., Reilly, R.: Automated processing of shoeprint images based on the fourier transform for use in forensic science. Trans. Pattern Anal. Mach. Intell. **27**, pp. 341–350 (2005)
9. Su, H., Crookes, D., Bouridane, A., Gueham, M.: Local image features for shoeprint image retrieval. In: British Machine Vision Conference 2007. (2007)
10. Zhang, L., Allinson, N.: Automatic shoeprint retrieval system for use in forensic investigations. In: 5th Annual UK Workshop on Computational Intelligence (2005)
11. Everingham, M., Zisserman, A., Williams, C.K.I., Van Gool, L.: The PASCAL Visual Object Classes Challenge 2006 (VOC2006) Results (2006)
12. Pavlou, M., Allinson, N.: Automatic extraction and classification of footwear patterns. In: Intelligent Data Engineering and Automated Learning, IDEAL 2006. pp. 721–728 (2006)
13. Su, H., Crookes, D., Bouridane, A.: Thresholding of noisy shoeprint images based on pixel context. Pattern Recognit. Lett. **28**, 301–307 (2007)
14. Su, H., Bouridane, A., Crookes, D.: Image quality measures for hierarchical decomposition of a shoeprint image. Forensic Sci. Int. **163**, 125–131 (2006)

Force Field Feature Extraction

The overall objective in defining feature space is to reduce the dimensionality of the original pattern space, while maintaining discriminatory power for classification. To meet this objective in the context of ear biometrics a novel force field transformation which treats the image as an array of mutually attracting particles that act as the source of a Gaussian force field has been developed. Underlying the force field there is a scalar potential energy field, which in the case of an ear takes the form of a smooth surface that

resembles a small mountain with a number of peaks joined by ridges. The peaks correspond to potential energy wells and to extend the analogy the ridges correspond to potential energy channels. Since the transform also happens to be invertible, and since the surface is otherwise smooth, information theory suggests that much of the information is transferred to these features, thus confirming their efficacy. Force field feature extraction, using an algorithm similar to gradient descent, exploits the directional properties of the force field to automatically locate these channels and wells, which then forms the basis of the characteristic ear features.

► Physical Analogies for Ear Recognition

Force Field Transform

An invertible linear transform which transforms an image into a force field by pretending that pixels have a mutual attraction proportional to their intensities and inversely to the square of the distance between them rather like Newton's Law of Universal Gravitation. Each pixel is assumed to generate a spherically symmetrical force field so that the total force $\mathbf{F}(\mathbf{r}_j)$ exerted on a pixel of unit intensity at the pixel location with position vector $\mathbf{r}_j$ by a remote pixel with position vector $\mathbf{r}_i$ and pixel intensities $P(\mathbf{r}_i)$ is given by the vector summation,

$$\mathbf{F}(\mathbf{r}_j) = \sum_i \left\{ \begin{array}{l} P(\mathbf{r}_i) \dfrac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \forall i \neq j \\ 0 \forall i = j \end{array} \right\}. \quad (1)$$

To calculate the force field for the entire image, this equation should be applied at every pixel position in the image. In practice this computation would be done in the frequency domain using Eq. 2 where $\Im$ stands for FFT and $\Im^{-1}$ stands for inverse FFT.

$$\text{forcefield} = \sqrt{M \times N}\,\Im^{-1}[\Im(\text{unitforcefield}) \times \Im(\text{image})]. \quad (2)$$

► Physical Analogies for Ear Recognition

Forensic

Forensic is the use of science or technology in the investigation and establishment of facts or evidence in the court of law.

► Skull, Forensic Evidence of

Forensic Anthropology

Forensic anthropology is the application of physical anthropology in special cases with forensic importance, such as to identify the skeletonized human remains.

► Skull, Forensic Evidence of

Forensic Applications, Overview

Christophe Champod
Institut de Police Scientifique, Ecole des Sciences Criminelles, Université de Lausanne, Switzerland

Introduction

The use of biometric data is a decisive process in ► forensic science that helps to establish a person's identity or associate two unknown persons. Forensic scientists realized that physiological or behavioral data could help to inform about, sort, and potentially individualize the persons involved in criminal offences. It is the case when (1) an unknown individual (living or his/her remains) has to be identified, (2) when biometric traces left by unknown individuals during activities of interest have to be traced back to their sources, or (3) when biometric traces have to be linked together in a series. Situations (1) and (2) require comparison between biometric information gathered from unknown sources and material of known (or declared as

such) origin, either on a one-to-one or on a one-to-many basis. In the latter case, data of known origin are organized in a database, allowing one-to-many searches. The third situation (3) compares biometric material from unknown sources and groups them according to potential (yet unidentified) sources. This last activity may involve the use of a database or may be carried out on a case-by-case basis.

Both situations (2) and (3) take advantage of what is known in forensic science as "Locard's exchange principle". Locard suggested in 1920 that forensic scientists can take advantage of the traces (such as fibers, paint, firearms discharge residues, dust, bloodstains, etc.) and marks (e.g., finger marks, footwear marks, toolmarks, etc.) exchanged between actors (e.g., a victim and an offender), and associated objects or scenes involved in criminal activities. The systematic search (potentially helped with detection techniques), preservation and analysis of these marks and traces will help to establish the relation between the actors, objects or scenes, and reconstruct the course of activities. The biometric features that can be helpful are many (and tend to increase in our modern society) and may consist of the following types of traces or marks:

1. Handwritten notes, including disputed signatures.
2. Finger marks and, by extension, any marks of friction ridge skin.
3. Barefoot impressions.
4. Bloodstain, semen stain, saliva stain, and any other biological fluids.
5. Earmarks.
6. Impression left by a face on airbags in car crashes [1].
7. Images of individuals (including images showing face, ears, or other identifying features) in stills or videos taken either from CCTV systems or any image-recording device.
8. Recording of a voice utterance of an individual (either in analog or digital form).

Biometric features have been used in forensic science for many centuries for some attributes (such as handwriting) and more recently for others (e.g., DNA profiling or recognition of faces from CCTV surveillance camera). However, in many respects, the application of biometry in forensic science is in contrast to the deployment of biometric systems in other areas (such as access control) this is mainly due to the unpredictable nature in terms of quality and quantity of the biometric features available to the forensic scientist, especially when they are collected as traces. These contrasts have been detailed elsewhere [2].

This overview is intended to cover the standard use of biometric features by forensic experts, either manually or with semi-automated or automated systems. While distinguishing the types of biometry helps to structure the text, it also helps to distinguish between the investigative and evaluative use of the techniques.

In the investigative mode, marks and traces are questioned to provide leads that may help to focus the inquiry, without making any reference to a potential source at the outset. Typical questions are as follows:

1. Can we link these criminal incidents the biometric evidence recovered?
2. On the basis of the recovered material, can we make any inference regarding the sex, age, ethnicity, or other physical attributes of the donor?
3. By comparing this biometric entry with a forensic database, is it possible to prepare a short list of potential donors?

In the evaluative mode, the questions are directed towards a source that is available, to provide control material to be compared against the unknown material. The case is then focused on one or more identified persons, with a view to help assess whether they are the source of the recovered material.

The references used hereinafter are limited to major textbooks or papers for each evidence type considered. Some evidence types are largely covered in specific chapters of this encyclopedia and address specifically the issues associated with automatic recognition.

Admissibility of Biometric Evidence in Court

The question of admissibility is crucial when the scientific element is used in the evaluative mode, especially when it is expected to find and present identification evidence in a courtroom. In the United States of America, the criterion for admissibility was traditionally based on the *Frye* rule (1923), which invites the judge – acting as a gatekeeper – to assess if the

technique has gained general acceptance within the relevant scientific community. It was under *Frye* that numerous human identification evidence types such as fingerprints, handwriting evidence, and DNA gained acceptance. Earprint, however, failed to pass the *Frye* test. The *Frye* standard has been revised following a ruling of the US Supreme court in *Daubert v. Merrell Dow Pharmaceuticals* (1993) and its progeny [3]. *Daubert* gave an interpretation of the Federal Rule of Evidence (FRE) and required the judge, still acting as gatekeeper, to assess more than only the general acceptance and include five criteria:

1. Whether the expert's technique or theory can be or has been tested, that is, whether the expert's theory can be challenged in some objective sense.
2. Whether the technique or theory has been subject to peer review publication.
3. The known or potential rate of error of the technique or theory when applied.
4. The existence and maintenance of standards and controls.
5. Whether the technique or theory has been generally accepted in the scientific community.

Nowadays in the United States of America, at the federal level, *Daubert* is in force. States may apply either *Frye* (or similar state decisions) or *Daubert*.

Daubert led to an increased number of challenges in court and forced the forensic community to articulate in detail the foundations of their disciplines, even in areas that had gained acceptance in US courtrooms for numerous years (see http://www.daubertontheweb.com/).

In Europe, there is no specific admissibility rule regarding scientific evidence. The principle of the judges' free evaluation of the evidence prevails. Hence, it is not surprising to see currently, a limited debate in the European jurisprudence regarding the admissibility of identification evidence.

From Anthropometry to Fingerprinting and AFIS

In the face of the absence of any reliable means of identifying recidivists, Alphonse Bertillon proposed in 1881 a classification and retrieval method based on anthropological measures. He took advantage of the fact that bone lengths remain constant in adulthood. It varies from individual to individual and can be measured with reasonable precision. Eleven precise measurements (height of the individual, length of outstretched arms, height of trunk, length and width of the head, length of the left middle finger, the left foot, left forearm, and right ear) combined with a mention of the color of the iris, were proposed to establish an anthropometric form for each arrested individual. This anthropometric record was completed with a photograph of the face and a standardized description of particular marks that can be filled and retrieved. The system was essentially used as an investigative measure to help identify individuals arrested multiple times (in time and place). The combination of anthropometric measurement, forensic photograph, and the standard description of the face was coined "Bertillonnage". A rapid spread of Bertillonnage has been observed at the turn of the 20th century across the world police departments and penal institutions [4]. The limitations of this technique were quickly noticed: (1) uneven distributions of the measurements; (2) correlation between features; (3) inter-operator variations, and (4) the imperative need of the body of the individual because no anthropometric traces are left on crime scenes (at the time).

In 1880, Herschel, a colonial administrator in India, published his proposal to use finger prints to identify individuals. At the same time (in 1880), Faulds, a medical missionary in Japan, proposed using finger prints for investigative identification purposes as well, as finger marks could be detected on crime scenes. Fingerprinting became a credible alternative to anthropometry for identification of habitual offenders, when Galton presented in 1892 the basic axioms of fingerprinting, including permanence (based on Herschel's work and data), discriminative power (Galton published the first statistical model on the fingerprint variability), and the possibility of reliably classifying fingerprints into three basic patterns. The classification method was then greatly improved by Henry and gained a large acceptance in English-speaking countries. Almost simultaneously, Vucetich , proposed a simpler system, based on Galton's initial proposal, that proved very successful for small to medium size databases. From the early 1900s, fingerprinting became the sole means of identification of habitual offenders throughout the world [5]. The use here is investigative (search for a potential candidate in a repository of fingerprint forms) and evaluative (verification on a

one-to-one basis), based on ten print records (taken from living or dead individuals).

An additional benefit of papillary ridge skin is that marks are left on the crime scene and are either readily visible or detectable using adequate detection techniques [6]. The first cases of identification of criminals through the finger marks left by them are attributed to Vucetich (1892) and Bertillon (1902). These marks can be searched against a fingerprint file or, of late, in an AFIS (Automatic Fingerprint Identification System). With the increase of tenprint cards and the difficulty of searches based on papillary marks, research in automatic retrieval processing systems took off in parallel with the technological advances since the 1960s. All forensic AFIS are nowadays largely based on minutiae matching both for finger and palm impressions [7]. The main advantage of an AFIS is the ability to compare a single print or mark, as well as a tenprint card, to the whole database. Note that AFIS provides a list of best candidates (according to a scoring/ranking metric). The identification process is not carried out by the system, but processed manually by an expert (through a dedicated user interface) in exactly the same way as if the potential candidate prints were suggested as a result of the usual police inquiry.

Other Biometric Characteristics used for Human Identification Purposes

The use of *deoxyribonucleic acid (DNA)*, a chain of nucleotides contained in the nucleus of our cells, has been a major breakthrough in forensic science to help in the identification of unknown individuals or biological samples left by them. Nuclear DNA can be extracted from all biological tissues (blood, saliva, urine or semen, from hair (with roots) and skin cells left by contact with the skin). For identification purposes (of the living or dead), samples are obtained from blood, saliva, or bones. The most common analysis of nuclear DNA is focused on STRs (short tandem repeats) [8]. STRs are repetitive sequences at a given location of the DNA molecule, of non-coding nature, which show a large and well-documented polymorphism. At a given locus, one individual will show two specific numbers of repetitions of the given sequence of nucleotides. These two numbers called *alleles* then give the biometric template for that locus. Note that one allele results from the genetic transmission from the biological father, and the other from the biological mother. The constitution of DNA databases to assist investigation is simple and has been developed in most countries.

Forensic applications of DNA profiling for human identification are numerous for STR analysis and cover (1) the comparison of an unknown profile (for example, from human remains) of an individual to a database of known profiles or profiles of potential relatives; (2) filiation testing when putative genitors are available or alternatively with ancestors or descendants. The first introduction of DNA profiling in forensic science dates back to 1986 (the Pitchfork case in the UK), but the large development of practices started in the 1990s. Before that time, biological fluids were analyzed using blood grouping determination or analysis of various proteins or enzymes [9]. Most of these forensic analyses have been abandoned for identification purposes in favor of DNA profiling because of the limited sensitivity and discriminating power of these systems.

When nuclear DNA cannot be analyzed (typically because of the degradation of DNA), mitochondrial DNA (contained in the mitochondria and inherited through maternal lines) can be used. Its discriminating power, however, is much lower than STR nuclear DNA analysis.

Dental features are mainly used in the identification of human remains in cases of missing persons or mass disasters [10]. The features used range from the standard dental record (indication of missing teeth, restorations, crowns, etc.) to dental radiographs (tooth contours, relative positions of neighboring teeth, and shapes of the dental work). These anatomical features have shown very good stability and variability and the teeth serve as a suitable repository of manmade operations that will leave various marks and shapes. Alphanumerical data can easily be organized in databases and such systems are used operationally in cases of mass disasters (http://www.interpol.int/Public/DisasterVictim/Default.asp).

Forensic analysis of soft tissues can help in the identification of remains. Analysis of scars, incised wounds, burn marks, trauma, and medical/surgical intervention are typical either in the investigative or evaluative mode [11]. When no tissue is left, *forensic anthropology* becomes an essential part of forensic and archeological investigations [11, 12]. Following the recovery of unidentified skeletal remains, the forensic anthropologist

can assist in guiding the investigation to identify the sex, ethnic origin, stature, and age (and if it is a woman's remains, whether she had gave birth) of the deceased. This information is investigative in nature. The same applies to cranio-facial reconstruction from the skull to help in the search for a deceased person [13]. When reference material is made available (X-ray images from ante mortem medical documentation), its comparison with post mortem data can help to establish identity through the analysis of morphology, fractures, medical interventions on the bones, and frontal sinus shapes [14]. Most of these areas have not been subjected to extensive automation research [15].

Other Biometric Marks Left Following Activities of Forensic Interest

DNA profiles can be obtained from the marks left behind by activities of forensic interest, typically from stains of blood, saliva, urine or semen, and from hair (with roots) and skin cells left by mere contact. Extracts are amplified using a sensitive and selective DNA replication method known as *Polymerase Chain Reaction* (PCR). In practice sensitivity to levels below 100 pg of DNA (a few cells) can be achieved. Such sensitivity widens the investigative possibilities, allowing the analysis of biological stains of very limited quantity. These profiles can be used to compare a DNA profile obtained from biological material against profiles from known individuals. If a correspondence is obtained, then this information can be used as evidence in court. It is important to stress that a match between two DNA profiles does not establish conclusively an identification of sources. Indeed, although the selectivity of DNA profiling is very high, there exists a probability of random association. In addition, DNA analysis offers some investigative capabilities gathered through the systematic comparison of DNA profiles coming from various scenes or familial searches against the DNA database. Another investigative aspect is the use of specific DNA analysis (SNP for single nucleotide polymorphism) to infer (within defined uncertainty boundaries) iris or hair color, skin pigmentation, and ancestry background [16].

The morphology of the *ear* was considered by Bertillon as the most identifying part of an individual. This modality was thus quickly used for identification purposes in forensic cases, either on photographs (or still images from video recording [17]) or on *ear marks* left on crime scenes, for instance, on doors. Forensic ear or ear print comparison is traditionally completed by skilled examiners according to published principles and protocols [18]. It is important to note that there is a big gap in terms of quality between a well-taken photograph of a ear and its impression on a door; hence, the strength of the evidence may vary from case to case as a function of the quality and the extent of the available material. Ear print examination can be used in the investigative phase to constitute a series based on the collected marks or to estimate the height of the donor, or in an evaluative manner, to associate a recovered mark with the ears of a designated individual.

Barefoot impressions can be left either on crime scenes or inside the shoes [19]. In the first instance, their investigation will help to assess the sequence of events and associate or exclude a given individual from being at the source of these marks. In the second, the analysis can help to assess whether a given individual is the habitual wearer of the shoe. Their use in criminal investigations predates the use of finger marks [20]. Barefoot impressions have shown a very high discrimination power and allow, when the quality of the mark is adequate, to bring powerful evidence of the identity of the sources in court.

Bite marks can be left on various substrates (the skin of a victim, some food, etc.) and can be compared against the control material from potential donors. A full account of their detection and analysis can be found in [21]. If bite mark analysis is to continue to play a role in the judicial process, there is an urgent need for high quality studies that meet the levels of forensic and scientific scrutiny applied to the other disciplines within the criminal justice system [22, 23].

From time to time, *lip marks* can be recovered from objects that came into contact with lips. Forensic lip print analysis is a very anecdotal area.

Handwriting and signature are biometric attributes with a long history in forensic science. The principles and procedures used by forensic experts to assign questionable handwritten documents to known individuals are described in [24]. The forensic expert tries to assess existing similarities and dissimilarities between control and recovered samples through a subjective estimation of the individuality and variability of the material at hand. At the moment, the automatic techniques used for handwriting and signature recognition are in their

infancy, especially in forensic science. Following the *Daubert* challenges, the field has been the focus of an increased scrutiny as to its scientific underpinning. It has led to a new body of research that shows the fertile avenues of collaboration between biometric computer science and forensic science [25–27].

Analog/Digital Biometric Information Recorded in Investigations

Forensic speaker recognition can be defined as any process using speech signals to determine if a specific individual was the speaker of a specific declaration. Experts may reach opinions from a variety of techniques used alone or in combination: auditive comparison, visual comparison of spectrograms, and semi-automatic methods for extraction of specific parameters (e.g., formant frequencies). Auditive comparisons are more likely to be conducted by phoneticians. They assess voice characteristics (voice, speech, language, and linguistic) either subjectively or objectively (using signal processing tools). The visual spectrographic approach was first proposed in 1962. In 1976, the US National Academy of Sciences recommended the use of this approach in forensic cases cautiously [28]. Automatic speaker recognition is also used in forensic science (see related entry in this Encyclopedia). Several characterization and modeling tools have been developed for automatic speaker recognition. All are sensitive to voice modification in recording and transmission conditions and their performance worsens when the conditions deteriorate. In forensic cases, the recording conditions of the trace and the reference materials are rarely similar or ideal, but rather record in different and unconstrained conditions, i.e., through mobile communications (GSM) transmission and with background noise. Due to these factors, the comparison is often undertaken under adverse conditions.

Facial images are more and more available for forensic investigations. Forensic face recognition is generally carried out by dedicated experts using approaches based either on morphological analysis of facial structures, anthropometric measurements, or image superimposition [29]. The morphological approach is based on a nomenclature for the description of the physiological aspects of the nose, the forehead, and the ear. Additional information, such as facial wrinkles and scars, can also be used. As the description is rather subjective, variations between operators are observed. In addition, the features of the same individual change due to expression changes, photographic angles or aging, and the demonstration of their statistical independence is often weak. The anthropometric approach can be described as the quantification of physiological proportions between specific facial landmarks. This method is only used for the comparison of faces with the same orientation. In order to avoid any scale and absolute size differences between photographs, ratios are calculated from these landmarks. Lighting conditions, camera distortions, camera positioning, facial orientation, facial expressions, and aging may impact the measures. The superimposition-based approach is the juxtaposition or the superimposition of facial images, taken under the same acquisition conditions (the orientation, the pose, and the size).

These three main comparison approaches do not yet consider automatic face recognition (a subject covered in several chapters of this Encyclopedia). Automatic face recognition systems have a large role to play in the future, not only in dealing with the face as such but also taking advantage of lips [30] or other features. But before introducing any automatic face recognition in court, a full and systematic assessment of the system should be conducted under realistic conditions, using fit-for-purpose forensic efficiency measures.

The prevalence of images or videos in modern society opens the route to the development of new types of forensic biometry (some are already covered in this Encyclopedia, e.g. gait analysis). Is it not rare to observe anatomical features on images that can help towards the identification of the individual captured on these images? These features can be *skin details, scars, veins and tattoos*. Biometric developments are still in their early stages [31].

References

1. Yamazaki, K., Imaizumi, K., Kubota, S., Atsuchi, M., Noguchi, K., Yosino, M.: Experimental study on personal identification from faceprint on vehicle's airbag. Japanese Journal of Science and Technology of Identification. **9**(1), 19–27 (2004)
2. Dessimoz, D., Champod C.: Linkages between biometrics and forensic science. In: Flynn, P.J, Jain, A.K, Ross, A. (eds.) Handbook of Biometrics, pp. 425–459. Springer, New York (2007)
3. Saks, M.J., Faigman, D.L.: Expert evidence after Daubert. Annu. Rev. Law Soc. Sci. **1**(1), 105–130 (2005)

4. Cole, S.: Suspect identities: A history of fingerprinting and criminal identification. Harvard University Press, Cambridge, MA (2001)
5. Berry, J., Stoney, D.A.: The history and development of fingerprinting. In: Lee, H.C., Gaensslen, R.E. (eds.) Advances in Fingerprint Technology, 2nd edn. pp. 1–40. CRC Press, Boca Raton, FL (2001)
6. Champod, C., Lennard, C.J., Margot, P.A., Stoilovic, M.: Fingerprints and other Ridge Skin Impressions. CRC Press, Boca Raton, FL (2004)
7. Komarinski, P.: Automated fingerprint identification systems (AFIS). Elsevier, New York (2005)
8. Butler, J.M.: Forensic DNA typing, 2nd edn. Elsevier, Burlington, MA (2005)
9. Gaensslen, R.E.: Sourcebook in Forensic Serology, Immunology, and Biochemistry. US Department of Justice, National Institute of Justice, US Printing Office, Washington, DC (1983)
10. Sweet, D., Pretty, I.A.: A look at forensic dentistry – Part 1: The role of teeth in the determination of human identity. Br. Dent. J. **190**(7), 359–366 (2001)
11. Black, S.M. (ed.): Forensic Human Identification: An Introduction. CRC Press, Boca Raton, FL (2006)
12. Pickering, R.B., Bachman, D.C.: The Use of Forensic Anthropology. CRC Press, Boca Raton, FL (2000)
13. Iscan, M.Y., Helmer, R.P. (ed.): Forensic Analysis of the Skull: Craniofacial Analysis, Reconstruction, and Identification. Wiley, New York (1993)
14. Christensen, A.M.: Assessing the variation in individual frontal sinus outlines. Am. J. Phys. Anthropol. **127**(3), 291–295 (2005)
15. Falguera, J.R., Falguera, F.P.S., Marana, A.N.: Frontal sinus recognition for human identification. In: Vijaya Kumar, B.V.K., Prabhakar, S., Ross, A.A. (eds.) Biometric Technology for Human Identification V. In: Proceedings of the SPIE; 2008 March 18, 2008; Orlando, FL. SPIE; 2008. p. 69440S–9
16. Frudakis, T.: Molecular photofitting: Predicting Ancestry and Phenotype using DNA. Academic Press, Burlington, MA (2008)
17. Hoogstrate, A.J., van den Heuvel, C., Huyben, E.: Ear identification based on surveillance camera images. Sci. Justice. **41**(3), 167–172 (2001)
18. van der Lugt, C.: Earprint Identification. Elsevier Bedrijfsinformatie, Gravenhage (2001)
19. Kennedy, R.B., Yamashita, A.B.: Barefoot morphology comparison: A summary. J. Forensic Ident. **57**(3), 383–413 (2007)
20. Caussé, S.: Des empreintes sanglantes des pieds, et de leur mode de mensuration. Annales d'hygiène publique et de médecine légale. 1854;1 (2ème série):175–89
21. Dorion, B.J. (ed.): Bitemark Evidence. Marcel Dekker, New York (2005)
22. Pretty, I.A.: The barriers to achieving an evidence base for bitemark analysis. Forensic Sci. Int. **159**(Suppl 1), S110–S20 (2006)
23. Bowers, C.M.: Problem-based analysis of bitemark misidentifications: The role of DNA. Forensic Sci. Int. **159**(Suppl 1), S104–S9 (2006)
24. Huber, R.A., Headrick, A.M.: Handwriting Identification: Facts and Fundamentals. CRC Press, Boca Raton, FL (1999)
25. Marquis, R., Schmittbuhl, M., Bozza, S., Taroni, F.: Quantitative characterization of morphological polymorphism of handwritten characters loops. Forensic Sci. Int. **164**, 211–220 (2006)
26. Schomaker, L.: Advances in writer identification and verification. In: Ninth International Conference on Document Analysis and Recognition – ICDAR 2007, pp. 1268–1273 (2007)
27. Srihari, S., Huang, C., Srinivasan, H.: On the discriminability of the handwriting of twins. J. Forensic Sci. **53**(2), 430–446 (2008)
28. Bolt, R.H., Cooper, F.S., Green, D.M., Hamlet, S.L., McKnight, J.G., Pickett, J.M. et al.: On the Theory and Practice of Voice Identification. National Research Council, National Academy of Sciences, Washington, DC (1979)
29. Iscan, M.Y.: Introduction of techniques for photographic comparison: Potential and problems. In: Iscan, M.Y., Helmer, R.P. (eds.) Forensic Analysis of the Skull, pp. 57–70. Wiley-Liss, Inc., New York (1993)
30. Chorań, M.: Human lips as emerging biometrics modality. In: Image Analysis and Recognition: 5th International Conference, ICIAR 2008, Póvoa de Varzim, Portugal, June 25–27, 2008 Proceedings, p. 993–1002. Springer, Berlin (2008)
31. Jain, A., Lee, J-E., Jin, R.: Tattoo-ID: Automatic tattoo image retrieval for suspect and victim identification. In: Advances in Multimedia Information Processing – PCM 2007, pp. 256–265 (2007)

Forensic Barefoot Comparisons

Brian A. Yamashita[1], Robert B. Kennedy[2]
[1]Forensic Identification Operations Support Services, National Services and Research, Royal Canadian Mounted Police, Ottawa, ON, Canada
[2]Royal Canadian Mounted Police (retired), Ottawa, ON, Canada

Synonyms

Barefoot morphology comparison; Footprint comparison

Definitions

Forensic barefoot comparison, or barefoot morphology comparison, describes the comparison of impressions of the weight-bearing areas of feet in an attempt to include or exclude a suspect as someone linked to a crime scene. A bare or socked foot impression found at the crime scene can be compared to inked barefoot

impressions and footprint casts taken from a suspect. Similarly, a link to footwear matched to a crime scene can be determined by comparing the insoles of the crime scene footwear to footwear seized from a suspect, or to inked impressions and casts taken from a suspect.

Introduction

Barefoot morphology comparison refers to the examination of the weight-bearing areas on the bottom of a human foot, when ridge detail is not present, to establish a link between the bare foot of an individual and a footprint impression found at a crime scene [1–3]. In the case of footwear linked to a crime scene, comparison can be made to shoes seized from a suspect, or to inked impressions or casts taken from a suspect. Research has indicated that the shapes of footprints are sufficiently variable to make it possible to include (as having possibly made the impression) or exclude (as definitely not having made the impression) a suspect as being the person who created a particular footprint at a crime scene [3]. As an example, Fig. 1 show barefoot impressions taken from identical twins, illustrating that even twins can be differentiated based on their footprints.

Forensic Barefoot Comparisons. Figure 1 Barefoot impressions taken from identical twins, illustrating the variability of footprints, even for twins.

When a crime scene is being examined, it is common to find footprints that might be those of the perpetrator of the crime. If ridge detail is developed in a barefoot impression, the comparison to a suspect's foot can be carried out in exactly the same fashion as a fingerprint comparison [4, 5]. If enough ridge detail, with sufficient clarity, is available for comparison, a positive identification may be forthcoming. However, if the barefoot impression is smudged or unclear for any other reason, or if it is a socked impression, then recourse can be made to barefoot morphology comparison.

Similarly, when a shoe has been positively identified back to a crime scene, and no suspect has been found in possession of the footwear in question, recourse can again be made to barefoot morphology comparison. This can be accomplished by comparing the impressions on the insole inside the crime scene shoe to the impressions inside a similar shoe worn by the suspect, or to inked impressions and casts seized from the suspect.

The comparison itself is similar to a toolmark or tire track comparison, where the foot has acted like a tool or a tire in creating an impression at the crime scene. The shapes of various parts of the foot are compared to see if there is correspondence between the crime scene impression and the suspect exemplars. In a footwear example, the impression to be compared has been made on the insole of the identified footwear.

Background

Although footprints in general look quite similar, it has long been assumed that careful examination of barefoot impressions could be used to differentiate between people. Historically, in various societies, trackers have been trained to be able to pick up someone's trail and to follow the person based on their footprints [3].

Footprint evidence was presented in court as early as the late nineteenth century, when a criminal was convicted in 1888 based on his footprint. Other cases have since been documented in the forensic literature, mainly from Europe and North America [6]. Much of this early casework was based on the assumption that

footprints were unique to the individual, without many studies to support this hypothesis.

Whenever a new means of including or excluding suspects is being introduced in court, the basis for the comparison must be justified. Some early work on footprint variability was carried out in India [7], while in North America, Dr. Louise Robbins, an anthropologist, carried out studies on the individuality of footprints in the 1970s [8]. In the 1980s, the Federal Bureau of Investigation (FBI) collected and compared footprint impressions from hundreds of volunteers to show how variable barefoot impressions might be [9, 10].

The Royal Canadian Mounted Police (RCMP) began research in this area in the 1990s [11]. Inked barefoot impressions were collected from thousands of volunteers for entry into a computerized database. As each footprint was measured and entered, it was compared with previous impressions in the database to ensure that another foot did not share the same measurements. A statistical analysis of the impressions was carried out to illustrate how variable barefoot impressions are [12]. Even with a limited number of samples and measurements, probabilities on the order of 1 in a billion were achieved.

Collecting Evidence [3]

When a bare or socked foot impression is found at a crime scene, the investigator must document the evidence correctly. Photographs, with a scale included, should be taken. If required, barefoot impressions can sometimes be enhanced *in situ* using fingerprint powder or chemical techniques, especially for impressions in blood. Impressions can then be lifted, or, depending on the surface, the entire impression can be removed from the scene. All of this evidence will have to be sent to the expert who is doing the final barefoot morphology comparison.

Similarly, when footwear impressions are found, they should be thoroughly documented. Again, enhancement techniques can be used to make the impressions more visible. If accidental characteristics are noted, there is the possibility of positively linking a shoe to the impression found at the crime scene.

If a suspect is arrested, his feet and his foot impressions must be well-documented. Several photographs, including a scale, should be taken of the feet to ensure that the tops, sides, and bottoms are all recorded. Foam impressions should be taken for later casting. Inked standing and walking impressions should also be obtained. Standing and walking impressions should also be obtained with the suspect wearing a pair of socks.

If the case involves footwear impressions, attempts should be made to seize similar footwear from the suspect. The best comparison would be of a shoe insole with another shoe insole. However, the feet of the suspect should also always be recorded in the same manner as described above.

The Comparison Process [3]

Barefoot morphology comparison should only be undertaken by an adequately-trained specialist. The RCMP has conducted barefoot comparison courses in North America and Europe, training forensic specialists from several countries. A group of doctors has recently formed a forensic podiatry sub-committee within the International Association for Identification (IAI), currently establishing its own criteria for training and standards.

As in any other physical comparison, class characteristics are compared first. The overall size of the foot and the number of toes making contact with the ground would be considered class characteristics, and can be used to quickly eliminate a suspect foot.

The shape and placement of the toes, the shape of the ▶ metatarsal ridge, the length and width of the arch, and the contour of the heel can be examined and compared (see Fig. 2). Any unexplained feature can be used to eliminate a suspect, while correspondence of features means that the suspect foot remains included as a possible source of the crime scene impression. Some examiners will make positive identifications based on foot morphology [13], while others will only go as far as a strong likelihood that the same foot made both the crime scene print and the exemplar [3]. Because the uniqueness of barefoot impressions has not been proven, and crime scene and inked impressions are not exactly reproducible, current RCMP policy advises examiners against making positive identifications. Positive identifications may be possible when flexion creases, marks, and scars are visible [14].

When the impression is three-dimensional, like a footprint in mud, it is important to try to obtain a

Forensic Barefoot Comparisons. **Figure 2** Comparing the toes and the contour of the metatarsal ridge.

three-dimensional replica of the suspect's foot for comparison. The use of foam impression material and dental casting material should be considered.

In the case of footwear that has been positively identified back to the crime scene, and where footwear has been seized from the suspect, the outer areas of the shoe can be examined to confirm that wear on the shoes is similar. The barefoot impressions made on the insoles can then be compared in much the same fashion as the barefoot comparisons described above. The inside uppers of the footwear can also be examined to look for agreement or disagreement of wear and damage. If footwear cannot be seized from the suspect, then the examiner must make-do with inked impressions and casts of the suspect's bare feet, bearing in mind the changes in morphology caused by the foot being constricted in the shoe.

In Court

Barefoot morphology comparison is not yet a routinely-accepted forensic technique. As such, it is still vigorously challenged when it is presented in court. In Canada, the testimony has been successfully defended in Mohan and Voir Dire hearings [15], (R v Dimitrov was overturned because the jury put too much emphasis on "could have been" testimony, and not because of the technique itself) while in the US it has withstood the scrutiny of "Rule 702," Frye, and Daubert challenges. An early case that was sent back on appeal when the judge felt that not enough background research had been done has been successfully re-tried in light of more recent published research. In essence, the courts have started to recognize the scientific foundation upon which the evidence is based.

Besides Canada and the United States, barefoot comparison testimony has been tendered in several countries around the world. In Israel, barefoot morphology comparison testimony was appealed all the way up to the highest court, and upheld. As the technique becomes more accepted, prosecutors and defense lawyers will soon start to look for opportunities where this type of evidence might be useful.

Conclusion

Barefoot morphology comparison refers to the comparison of the weight-bearing areas of feet in an effort to include or exclude a suspect as being linked to a crime scene. Background research has established the variability of barefoot impressions, justifying their use in a forensic context. Crime scene investigators should be aware of this technique and should always be looking for suitable evidence of this kind.

Related Entries

- ▶ Earprints
- ▶ Fingerprints

References

1. Bodziak, W.J.: Footwear Impression Evidence, pp. 381–411. CRC Press, Boca Raton, FL (2000)

2. Kennedy, R.B.: Bare footprint marks. In: Siegel, J.A., Saukko, P.J., Knupfer, G.C. (eds.) Encyclopedia of Forensic Sciences, pp. 1189–1195. Academic Press, London (2000)
3. Kennedy, R.B., Yamashita, A.B.: Barefoot morphology comparisons: a summary. J. Forensic Ident. **57**(3), 383–413 (2007)
4. Lemieux, M.: Histoire de Pied/a foot story. Identif. Can. **25**(4), 16–17 (2002)
5. Watkins, D., Brown, K.C.: The case of the toe print. J. Forensic Ident. **57**(6), 870–873 (2007)
6. McCafferty, J.D.: The shoe fits. The Police J. **28**(2), 135–139 (1955)
7. Puri, D.K.S.: Footprints. Int. Police Rev. **187**, 106–111 (1965)
8. Robbins, L.M.: The individuality of human footprints. J. Forensic Sci. **23**(4), 778–785 (1978)
9. Lovejoy, O.C.: Methods of Footprint Analysis. Seminar in Footprint and Shoeprint Identification. Federal Bureau of Investigation, 29 April 1984
10. Bodziak, W.J., Monson, K.L.: Discrimination of individuals by their footprints. Paper presented at 11th meeting of International Association of Forensic Sciences, Vancouver, BC (1987)
11. Kennedy, R.B.: Uniqueness of bare feet and its use as a possible means of identification. Forensic Sci. Int. **82**(1), 81–87 (1996)
12. Kennedy, R.B., Chen, S., Pressman, I.S., Yamashita, A.B., Pressman, A.E.: A large-scale statistical analysis of barefoot impressions. J. Forensic Sci. **50**(5), 1071–1080 (2005)
13. DiMaggio, J.A.: The Foot as a Forensic Tool. Paper presented at the 55th annual meeting of the American Academy of Forensic Sciences, Chicago, IL (2003)
14. Massey, S.L.: Persistence of creases of the foot and their value for forensic identification purposes. J. Forensic Ident. **54**(3), 296–315 (2004)
15. Richard, C.: Case law: [2002] Ontario superior court R. V. Arcuri. Identif. Can. **25**(4), 18–20 (2002)

Forensic DNA Evidence

T. Hicks, R. Coquoz
Institut de police scientifique, Ecole des sciences criminelles, Lausanne, Switzerland

Synonyms

DNA analysis; DNA profiling; DNA typing

Definition

Deoxyribonucleic acid (DNA) is a large molecule present in all living cells (e.g., animals, plants, viruses). As a tape allows the storage of a recording, DNA allows the storage of genetic information. It consists of two long chains of nucleotides twisted in a double helix. There are four types of nucleotides designed by the name of their bases: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). The genetic information is encoded in the sequence of nucleotides of the DNA molecule. Part of its name originates from its localization in the nuclei of the cell. However, the acronym is also used for ▶ mitochondrial DNA (mtDNA), which is the DNA present in the mitochondria of the cells. It is transmitted only by the mother.

Introduction

Although DNA profiling has high discriminating power and can help establishing the biological identity of a person, it is not used as a biometric yet. Indeed, the results of the analysis are not immediate (nor yet amenable to full automation), the cost of analysis is high, and there are contamination and transfer issues. Moreover, as parents transmit biological material to their children, DNA can be intrusive and reveal unknown family relationships.

Most of the text that follows is based on [1]. Another standard text in English is [2].

DNA: Basic Concepts

Each human cell contains biological information; each of the cells thus contains the same DNA. As DNA is a very long molecule (three billion base pairs), it is organised into 23 small bundles: the chromosomes. Each child receives two DNA: one from the mother and one from the father. Each person (and each cell from this person) has thus two copies of each chromosome, one from the mother and one from the father, giving a total of 46 chromosomes. They are numbered in pairs from 1 to 22, the 23rd pair being the sex chromosomes X and Y. On each of the chromosomes, there are genes (i.e., a zone where the DNA codes how to make a protein). The number of genes is estimated to be around 20,000–25,000. When reading DNA from one extreme to the other, one will encounter a code "START" to indicate that one can read from here how to make a given protein, and a code "STOP" to indicate that this is the end of the genetic information necessary for that protein. The code for the following protein does not begin immediately after the "STOP"

message and there are usually several thousands nucleotides in between that do not code genetic information. These are called noncoding DNA or junk DNA and represent 98% of the genetic material. The geographical distribution of the genes and the noncoding DNA is in principle identical across all individuals in a given species.

A geographical nomenclature has been derived to designate a given localization on the DNA strand of human chromosomes: the locus (pl. loci). For each locus, one copy is received from the father and one copy from the mother. These two copies are called alleles: therefore for each locus, a person will have two alleles. The set of alleles owned by a person for a locus is called his/her genotype. If the alleles transmitted by the father and the mother are the same, the individual is homozygote for this locus. If the two alleles transmitted are different, then the individual will be heterozygote at this locus. The loci are symbolized by a four-digit code, that is particularly useful for noncoding DNA. The first letter of the code is D (for DNA), the second is the chromosome number (1, . . ., 22, X, Y), the third element indicates the sequence type (S: a unique sequence; Z a sequence that has several copies, at different localizations, on the same chromosome; F a sequence that is part of a family with similar sequences that are encountered on several chromosomes); the last digits are a unique number that generally correspond to the order in which the sequences were discovered. As an example, the locus D18S51 is a DNA region on chromosome 18, with the serial number 51. This locus is commonly used in forensics because of its polymorphism and is present in different commercial kits (e.g., SGMplus ABI; Powerplex 16 Promega).

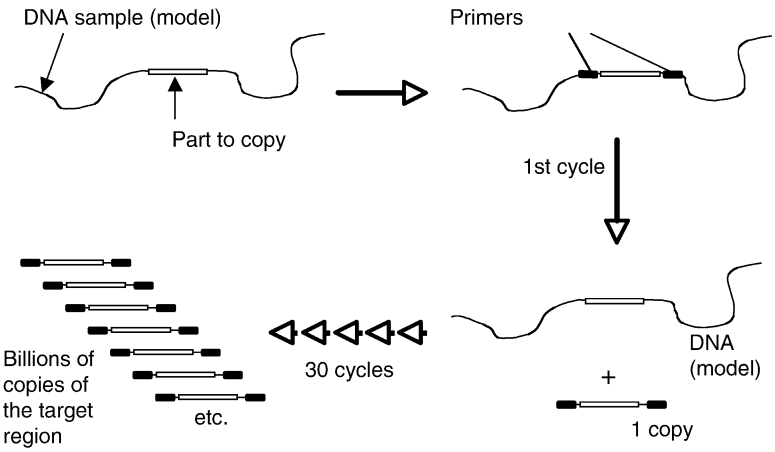
Since DNA is stable over the lifetime of an individual, and is reasonably resistant to chemical degradation, it is a good candidate for use in forensic science. One must further find a sensitive and not too expensive technique that enables the analysis of parts of the DNA that vary between individuals: the genetic markers. These markers should show as much as polymorphism as possible. There are two types of polymorphisms: sequence polymorphism (the nature of the nucleotides themselves differ from individual to individual, e.g., mtDNA, SNPs) and length polymorphism (the difference between individuals is based on the length of a given repetitive sequence of nucleotides, e.g., STR).

In 1985, thanks to the discovery of repetitive sequences, Sir Alec Jeffreys first applied the analysis of DNA to forensic science. As discussed previously, there are coding and non coding DNA. The repetitive sequences are zones in the DNA, where noncoding DNA seems to stutter. The human DNA has a very large number of different repetitive sequences (30% of the genome) and the length of the repetitive sequences varies between individuals. If the length of the repetitive sequence is larger than six nucleotides, one speaks of VNTR (Variable Number Tandem Repeats, Nakamura). If it is equal or smaller to six, one will speak of STR (Short Tandem Repeat). STRs have also been named microsatellites and VNTRs minisatellites. Nowadays, STRs are the standard targets of the routine analysis of forensic samples. Companies offer different type of kits allowing the analysis of several STRs at the same time (multiplex STR analysis). An example of some kits available is given below:

Forensic DNA Analysis

Before analysis, one has to sample the DNA on the crime scene and to sample the individual. As genetic information is theoretically the same in every cell, saliva is generally used for the later. On crime scenes, chemical tests to detect saliva, sperm, or blood can be used to help in finding the invisible stains. If a stain is detected, the object is either cut, or swabbed. As, it is possible to detect very low levels of DNA, it is highly recommended to wear gloves and a face mask when collecting DNA.

Once DNA has been collected, it will be extracted, purified, and quantified. Because there is often very little material in forensic samples, the specific zones of DNA that are the target of the analyses will be first amplified using a technique called Polymerase Chain Reaction (PCR). PCR is often compared to a DNA photocopier, where a given DNA segment is copied a given number of times. First, the zone to be amplified is delimited using two primers: one is placed at the beginning of the sequence and one at the end. During the copying cycles of the PCR process, copies of the original DNA zone will be produced. After one cycle, there will be the original DNA and one copy; after two cycles, there will be 2 more copies and after 30 cycles there will be about a billion copies (Fig. 1).



Forensic DNA Evidence. **Figure 1** From [1] representing the PCR Process.

Multiplex STR Analysis

PCR can be used to copy simultaneously several DNA fragments: one uses different primers, each corresponding to the DNA zone that has to be copied. This simultaneous amplification of different DNA zones is called Multiplex PCR. The PCR products are analyzed using Capillary Electrophoresis. This technique allows the separation of DNA fragments according to their sizes while they travel through a thin capillary. The large fragments move more slowly than small fragments. A peak is displayed on the results' graph, when the detector at the end of the capillary detects DNA molecules. The detector is able to differentiate between up to five different dyes. The detector can recognize STR alleles that have the same length but are labeled with different dyes.

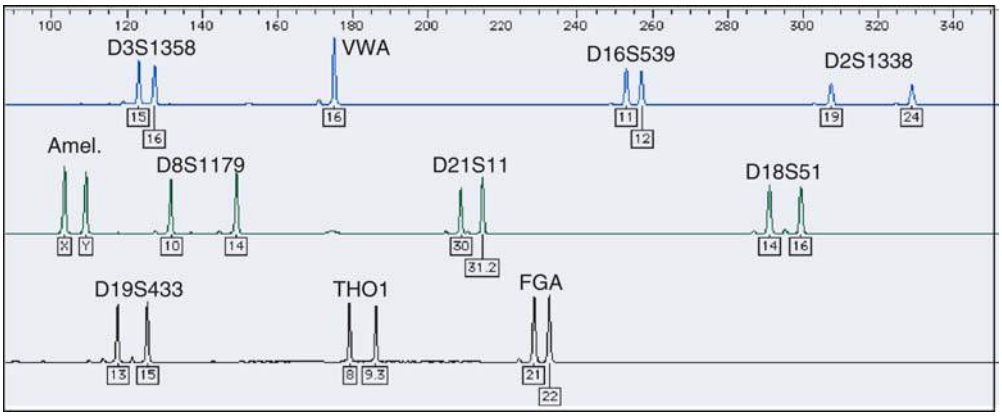
Theoretically, one can amplify dozens of different fragments in one operation, which saves considerable time. However, the design of multiplex STR analysis kits is not straightforward: the numerous primers involved must not interfere with each other; the set of alleles of one STR must be recognized from the alleles of the other STRs, either through the use of different labeling dyes or because they are in different size ranges. The different STR will be chosen if possible on different chromosomes, so that the transmission of the alleles from generation to generation can be considered independent.

PCR is unable to amplify efficiently large fragments. It is thus not possible to increase indefinitely the number of STR analyzed simultaneously with a multiplex kit. The practical limit is around 15–20 STRs. An example of a profile obtained with Multiplex STR analysis is shown below. The vast majority of forensic DNA analyses done today is multiplex STR analysis. It is the golden standard and fulfils most of the needs of forensic DNA analysis.

Figure 2 shows the result of an analysis performed on an individual with the SGMplus kit.

Nonautosomal DNA

An autosome is a nonsex chromosome. Nonautosomal DNA is a DNA that originates either from a sex chromosome (i.e., X, Y), or from mitochondria (mtDNA). Sex chromosome DNA is used to study the paternal heritage: if there is no mutation, a son will have the same Y chromosome than his brothers, his father, and his paternal lineage; a daughter will have one of her X chromosome identical to her father and her paternal lineage. Mitochondrial DNA is used to study the maternal heritage: a mother will transmit her mitochondrial DNA to her children, grandchildren, etc. These markers are useful for parentage testing. Since there are hundreds of copies of mtDNA per cell, mtDNA can also be used when the trace is degraded or/and there is very little nuclear DNA (e.g., hair without roots). STRY may also be useful in rape cases, when the male DNA of the agressor is not detected in the DNA profile because of the presence of vast amounts of female DNA in the mixture. When speaking of alleles observed for Y chromosome STR or mtDNA and STRX, one speaks of haplotype and not genotype, as a person owns only one single allele.



Forensic DNA Evidence. **Figure 2** Figure from [1] showing the result of an analysis performed on an individual with the SGM plus hit. The raw data are decomposed by color: the labels below the peaks show the designation of the detected alleles. She name above is the STR name. For most STRs, our individual has two peaks and is heterozygote for these makers. For VWA, this person precents only one peak and is therefore homozygote for that particular marker.

Low Template DNA Analysis

In essence, PCR allows a sensitivity at the single molecule level. However, for the standard DNA profiling, the potential of PCR is curbed to avoid the stochastic effects and contamination problems associated with an extreme sensitivity. The sensitivity of standard DNA profiling is thus adjusted to provide a DNA profile only when there is DNA from at least 50 cells in the evidence material. The concept of low template level analysis, or as previously known LCN (Low Copy Number), covers strategies designed to go beyond this limit. The usual strategy to reach such an extreme sensitivity is to increase the number of copying cycles of the PCR (e.g., 34 cycles instead of 28) [3]. Increasing sensitivity does not go without caveats: artefacts arise with alleles dropping in and dropping out. This added sensitivity also increases the issue of contamination. The use LCN profiles should induce very strict protocols to avoid contamination on the scene and at the laboratory. It requires cautious interpretation. In the Omagh Bombing, the use of low template level profiles was challenged, but was shown to be scientifically robust in the review conducted by Professor Caddy, Dr Linacre, and Dr Taylor [4].

SNPs and DNA Chips

SNPs (Single Nucleotide Polymorphisms) occur on average every 1,200 nucleotides: that means that depending on the individual the nucleotide in that position will differ. In theory there might be four variants (A, T, G, C) for the nucleotide; however, because of the reality of the evolution process only two variants are usually observed. The main caveat of SNPs is their limited sequence polymorphism. Multiplex analysis of a large number of SNPs can however overcome this limitation. SNPs can be analyzed using miniaturized devices called DNA chips. These are miniaturized systems allowing the analysis of hundreds of SNPs. The disadvantages of SNPs are their low polymorphism, and their very limited capacity to handle DNA mixture cases. SNP analysis certainly has good prospects for specific applications. ▶ Mitochondrial DNA polymorphisms are in essence SNPs. SNP analysis has better chances of success for highly degraded samples than STR analysis. Some specific SNPs have the capacity to provide morphological information (hair, eye colour, . . .) and Y chromosome SNPs should be able to provide useful ethnic information. Some companies (e.g., 23andme; Decode genetics; . . .) have started to offer to the public wide ranging SNP analyses providing information on their ancestry, and predispositions to possible diseases (see, for example, https://www.23andme.com/ or http://www.decode.com/). Although their approach is controversial and not forensically oriented, it is a good example of the power of the SNP analysis technologies. Applications of SNPs in context of forensic intelligence will be briefly mentioned later.

DNA Sequencing

This method consists in reading the sequence of the nucleotides of small pieces of DNA. It is the standard tool for mtDNA analysis. More recent "whole genome DNA sequencing" technologies are being developed. They are massively parallel sequencing approaches with the potential to provide the complete sequence of an individual in a single process. Their potential for forensic DNA analysis is not yet clear.

DNA used as Evidence and Interpretation

DNA (human, animal or vegetal) can be used in diverse areas: nature and species preservation, food control, missing persons, mass disaster identification, serious, and volume crime. As parents transmit part of their DNA to their children, parentage testing using DNA (autosomal and nonautosomal) is also common, whether for civil cases, historical cases (e.g., Thomas Jefferson, Nicolas II), immigration, or genealogy. DNA techniques are very sensitive and sources of DNA are diverse (e.g., blood, saliva, sperm, dandruff, skin, hair (nuclear DNA if root, mtDNA on shaft, bones, teeth...), which explains the great potential of forensic DNA analysis.

DNA evidence can help the court in assessing three types of propositions [5]: offence level propositions (e.g., the suspect has raped the victim vs. the suspect has not raped the victim), activity level propositions (e.g., the suspect has had sexual intercourse with the victim vs. the suspect has not had sexual intercourse with the victim) and source level propositions (e.g., the recovered DNA comes from the suspect vs. the recovered DNA originates from a person unrelated to the suspect). The higher the hierarchy (from source to offence level), the more information will be required to inform an opinion. Through DNA profiling and other means, the forensic scientist is usually only able to provide useful evidence for source and activity levels. If addressing source level proposition, the scientist will take into account the rarity of the DNA profile, estimating the match probability (Weir and Evett), [6]. When addressing activity level propositions, the forensic scientist will in addition take into account transfer and persistence of DNA (or/and blood pattern in the presence of blood, see [7]), as well as the relevance of the trace to the alleged activities.

DNA interpretation is a very large subject area [8, 6, 9–11] and (NRC reports I and II). With the large number of polymorphisms available, DNA of an individual can certainly be considered as unique. Thus, DNA profiles are frequently viewed as unique in the general public, but they are not. Only a limited number of polymorphisms are examined in DNA profiling, providing DNA profiles that are indeed very rare, with match probabilities smaller than 1 in a billion. The value of DNA evidence is usually assessed using the likelihood ratio approach (also called the Bayesian approach). The evidence is evaluated considering two propositions (e.g., the prosecutor's and the defence's): in the given example, the forensic scientist would, for example, assess the probability of the evidence given that the suspect has had sexual intercourse with the victim and the probability of the evidence given that the suspect denies knowing the victim. This approach insures an unbiased interpretation of the evidence.

DNA Used as an Intelligence Tool

With the launching of DNA databases in the 1990s, DNA has become a very useful intelligence tool. Most countries have national DNA databases or are in the process of doing so. The largest are the USA database and the England and Wales DNA database. Each country has its own legislation and own set of STR loci (in Europe there is a set a "core" loci that are used by all EU countries). The most common program used for storing and searching the DNA profiles is CODIS (Combined DNA Index System). The program was developed by the FBI and the private firm SAIC. The profiles are stored in different indexes (e.g., forensic profiles index, offender index, victim index, staff index, and missing person's relatives index). This allows comparing only profiles that should be (e.g., the profiles from relatives of missing persons will not be compared to crime scene profiles). Criteria for entering a profile in a database vary according to the country and the size of the database. DNA databases rely on the fact that the vast majority of crimes is committed by a small proportion of the population, that tends to re offend and on the fact that DNA profiles are extremely rare. This last prerequisite is not fulfilled with so-called partial DNA profiles (i.e., does not present results for all loci in the given kit because of DNA degradation), or with mixture DNA

Forensic DNA Evidence. Table 1 From [1] Table showing an example of multiplex kits available. Highlighted in dark gray are the markers chosen by the European Network of Forensic Science Institutes to warrant compatability across countries; in light and dark gray the markers chosen for the USA combined DNA Index System (CODIS) database

Supplier	ABI	ABI	ABI	ABI	ABI	ABI	Promega	Promega	Promega	Promega	Serac	Serac	Biotype
Kit Name	Cofiler	Profiler	Profiler Plus	SGM Plus	Identifiler	Sefiler	Powerplex 1	Powerplex 2	Powerplex 16	Powerplex ES	MPX2	MPX3	Mentype Nonaplex
Marker													
THO1	x	x		x	x	x	x	x	x	x	x	x	x
VWA		x	X	x	x	x	x	x	x	x	x	x	x
D21S11			X	x	x	x		x	x	x	x	x	x
FGA		x	X	x	x	x		x	x	x	x	x	x
D8S1179			X	x	x	x		x	x	x	x	x	x
D18S51			x	x	x	x		x	x	x	x	x	x
D3S1358	x	x	x	x	x	x		x	x	x	x	x	x
TPOX	x	x			x		x	x	x			x	
CSF1PO	x	x			x		x		x				
D13S317		x	x		x		x		x				
D7S820	x	x	x		x		x		x				
D5S818		x	x		x		x		x			x	
D16S539	x			x	x	x	x		x			x	
D2S1338				x	x	x						x	
D19S433				x	x	x						x	
Penta-D									x				
Penta-E								x	x				
SE33						x				x	x		x

profiles. That type of DNA profiles has to be used with more caution in conjunction of the database, especially, if the suspect population is large (i.e., the searched database is large). When there is no other forensic evidence than DNA to limit the suspect population, then the discriminating power of the technique must be higher, than if there are other means (e.g., partial fingerprints, modus operandi, micro-traces, and traditional police investigation information).

New approaches of the DNA database involve the use of partial profiles and familial searching for intelligence purpose. In general, to limit adventitious matches, partial profiles with less than six loci, for example, are generally not entered in the database, but they could be used to generate intelligence. Familial searching aims at helping the investigation when no match is found in the database. The technique consists in looking for profiles that share alleles with the crime scene profile. As it is more common for relatives to share part of their DNA than unrelated persons, there are examples where it was possible to find in the database a close relative of the offender. The Forensic Science Service (UK) has been able to solve a couple of famous cases using this method. There are ethical issues to consider when using this technique [12].

The analysis of some SNPs can predict physical characters (such as red hair, eye colour, . . .) based on the analysis of the crime scene sample (see www.dnaprint.com).

Conclusion

The advent of DNA analysis and DNA databases has revolutionised forensic science, police investigation and the whole criminal justice system. It is anticipated that automation will play tomorrow an even more important role than today. With the advent of ultra-sensitive methods, the relevance of the recovered material, the questions of transfer and persistence of DNA will become the core of interpretation. Today, the research in this area is still scarce and needs to be developed. Regarding the techniques, analysis of STR is here to stay for years. Other fascinating techniques for SNP analysis or even whole genome sequencing are coming. But they will not improve much the performance of DNA profiling and their use will require very difficult validations before they can diffuse widely into the routine practice. As it was portrayed in the film "Gattaca" by Mike Nichols, 1997, it seems clear that DNA on a chip and DNA as a biometric system at the finger tip will be available one day. But that day is not yet at the horizon.

Related Entries

► LCN DNA/Low Template Level

References

1. Coquoz, R., Taroni, F.: Preuve par l'ADN –la génétique au service de la justice. Presses Polytechniques et Universitaires Romandes, Lausanne (2006)
2. Butler, J.M.: Forensic DNA Typing. Biology, Technology, and Genetics of STR Markers. 2nd edn. Elsevier Academic, Burlington, MA (2005)
3. Gill, P., Whitaker, J., Flaxman, C., Brown, N., Buckleton, J.: In investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Foreusic Science International **112**(1), 17–40 (2000)
4. Caddy, B., Linacre, G.R., Taylor, A.M.T.: A Review of the Science of Low Template DNA Analysis. (2008) doi: http://police.homeoffice.gov.uk/publications/operational-policing/Review_of_Low_Template_DNA_1.pdf?view=Binary
5. Cook, R., Evett, I.W., Jackson, G., Jones, PJ., Lambert, J.A.: A hierarchy of propositions: deciding which level to address in casework. Science Justice **38**, 231–240 (1998)
6. Buckleton, J., Triggs, C., Walsh, S.J.: Forensic DNA Evidence Interpretation. CRC, Boca Raton (2005)
7. Bevel, T., Gardner, R.: Blood Pattern Analysis with an Introduction to Crime Scene Reconstruction, 3rd edn. CRC, New York (2008)
8. Evett, I.W., Weir, B.S.: Interpreting DNA Evidence – Statistical Genetics for Forensic Scientists. Sinauer, Sunderland (1998)
9. Aitken, C.G.G., Taroni, F.: Statistics and the Evaluation of Evidence for Forensic Scientists, 2nd edn. Wiley, Chichester (2004)
10. National Research Council, Committee on DNA Technology in Forensic Science, Board on Biology, Commission on Life Sciences. DNA Technology in Forensic Science, National Academy Press, Washington, D.C. (1992)
11. National Research Council, Committee on DNA Forensic Science. The Evaluation of Forensic DNA Evidence. National Academy, Washington, D.C. (1996)
12. Williams, R., Johnson, P.: 1Forensic DNA Databasing: A European Perspective. Interim Report. Unviersity of Durham, Durham (2005)
13. Robertson, B., Vignaux, GA.: Interpreting Evidence – Evaluating Forensic Science in the Courtroom. Wiley, Chichester (1995)

Forensic Evaluation of Fingerprints and Fingermarks

▶ Fingerprint, Forensic Evidence of

Forensic Identification Based on Dental Radiographs

▶ Dental Biometrics

Forensic Science

Forensic science refers to the applications of scientific principles and technical methods to the investigation of criminal activities, in order to establish the existence of a crime, to determine the identity of its author(s) and their modus operandi.

▶ Forensic Applications, Overview

Forensic Speaker Recognition

▶ Voice, Forensic Evidence of

Forgery Attempt

Active forgery attempt is an impostor attempt in which an individual tries to match the stored template of a different individual by presenting a simulated or reproduced biometric sample, or by intentionally modifying his or her own biometric characteristics.

▶ Influential Factors to Performance

Forgery Sign

Synonyms

Forgery signature; Impostor sign; Mimicked sign

Definition

Forgery sign is an illegal sign by simulating or tracing a genuine signature. There are two kinds of forgery signs such as "substitution or random" and "freehand or skilled." The former is called as "zero effort" forgery, because the forger uses his or her own signature instead of the signature to be tested. The later includes signatures imitated as closely as possible by simulating or tracing a genuine signature.

▶ Signature Matching

Forward-Backward Algorithm

The Forward–Backward algorithm is the conventional, recursive, efficient way to evaluate a Hidden Markov Model, that is, to compute the probability of an observation sequence given the model. This probability can be used to classify observation sequences in recognition applications.

▶ Hidden Markov Models

Fourier Transform

Mathematically, the continuous Fourier transform is one of the specific forms of the Fourier analysis.

It transforms the original function in the time-domain into another function in the frequency domain. The term "Fourier transform" can refer to either the frequency domain representation of a function or to the process/formula that transforms one function to another.

► Face Recognition, Component-Based
► Image Pattern Recognition
► Iris Encoding and Recognition Using Gabor Wavelets
► Iris Recognition Using Correlation Filters

Fovea

The fovea is a small depressed region at the center of the macula, the central area of the retina. There, the inner retinal layers are shifted aside, allowing light to pass unimpeded to the photoreceptors. Only tightly packed cones, and no rods, are present at the foveola, the center of the fovea. The elongated axons of these cone cell bodies are called Henle fibers. The fovea is the region of maximum visual acuity.

► Anatomy of Eyes

Fragile Bits

► Iris Template Extraction Via Bit Inconsistency and GRIT

Fraud Deterrence

► Fraud Reduction, Applications
► Fraud Reduction, Overview

Fraud Mitigation

► Fraud Reduction, Applications
► Fraud Reduction, Overview

Fraud Reduction, Applications

Victor Minchih Lee
International Biometric Group, New York, NY, USA

Synonyms

Biometric fraud reduction; Duplicate detection; Fraud deterrence; Fraud mitigation

Definition

Fraud is conventionally defined as the deliberate perversion or withholding of veracity in order to induce another to surrender something of value. In the context of biometrics, the item of value is typically an identity or a privilege associated with an identity.

For the purposes of this entry, fraud reduction in a biometric applications context refers to the use of biometric technology's duplicate detection capabilities to deter, inhibit, and mitigate fraud. Duplicate detection refers to the discovery of multiple identities claimed by a single, given individual.

Introduction

For numerous decades, individuals have sought to misrepresent their identities for the sake of obtaining benefits and privileges to which they are not properly entitled. Such fraud can be costly – both financially and politically. For fiscal year 2007/2008, the United Kingdom's Department for Work and Pensions estimated that it overpaid about £2.7 billion in housing-related benefits, alone, due to fraud and error [1]. From October 2002 to September 2005, the US Justice Department indicted 40 voters (21 noncitizens) for illegal voting or voter registration fraud [2].

In 2000, following the United States' presidential election, a study in Georgia, USA, discovered over 15,000 deceased individuals on the state's active voting rolls. The US Federal Election Commission also discovered 502,968 names on Alaska's 1998 voter rolls – yet only 437,000 eligible voters were estimated by the statewide census conducted that year [3]. In both cases, several thousand invalid, but influential, votes could have been cast in close elections by individuals assuming others' identities or fake identities.

In the past, such fraudulent actions were enabled by the tendency to ascertain identity based upon documentation with little – if any – connection to the distinctive characteristics of the legitimate document holder, aside from often replaceable or forgeable photographs. Authenticity of transactions was assured more by anti-forgery, document-oriented techniques (such as watermarks, holographs, security strips, microlines, intaglio printing, etc.) rather than by examination of the document bearer.

The advent of biometrics, however, has enabled a shift of focus from predominantly documents to a mix of documents and individuals. In 1858, the United Kingdom's William Herschel of the Civil Service of India was precocious in his decision to capture employee palmprints to help distinguish amongst his native Indian staff on paydays [4]. Today, automated biometric capture and processing systems allow for quick determinations and verifications of identity. They enable the detection of individuals who may assume multiple nominal identities through various documents, but who are really the same, single entity.

This duplicate detection capability deters and inhibits fraud in applications including:

1. Benefits issuance and disbursement
2. Voter registration
3. Visa shopping
4. Border control
5. Consumer recognition and
6. Time and attendance monitoring

This entry introduces and provides examples of the first four aforementioned applications. Consumer recognition (including check cashing) and time and attendance monitoring are both addressed in other entries.

Benefits Issuance and Disbursement

Governments are often responsible for the proper and equitable distribution of benefits to their qualified citizenry. With large populations of potential benefits recipients, however, it can be a logistics challenge to keep track of who is a qualified recipient and whether or not they have previously claimed a given benefit and are attempting illegitimately to reclaim the same benefit (a phenomenon sometimes referred to as "▶ double dipping").

Biometrics can help mitigate the problems associated with such challenges. Initially, when biometrics are first captured and associated with a given identity in an enrollment process, biometrics are particularly vulnerable and dependent on the legitimacy and robustness of ▶ breeder documents. But once a ▶ nominal identity has been paired with a biometric, a government can be relatively certain that whenever that biometric is presented, it is presented by the individual with that same nominal identity and not an imposter. This is because of the fundamental assumption and belief that biometrics based on individual physiological and behavioral characteristics are more difficult to steal and forge than are documents.

One example of the benefits that can ensue is the deployment of fingerprint recognition technology by the United States of America's Texas Health and Human Services Commission (HHSC). HHSC sought to ensure that Texas' limited Medicaid benefits be distributed only to the truly needy. HHSC wanted to make sure that it was paying for services actually rendered and delivered only to those authorized to receive Medicaid benefits. By leveraging fingerprint biometrics, HHSC sought to confirm that authorized Medicaid recipients were indeed physically at treatment facilities when Medicaid benefits were disbursed [5].

Another example of biometrics applied to fraud reduction in benefits issuance is the Andhra Pradesh Ration Card Entitlements program in India. In this deployment, first announced on 16 June 2005, iris recognition systems were employed in the issuance of food ration cards by the state of Andhra Pradesh [6]. Through the incorporation of biometrics in the program, Andhra Pradesh officials seek to deter its citizenry from selling or sharing their food ration cards, as well as returning to claim multiple cards under different nominal identities.

Voter Registration

The validity of elections critically depends on ensuring that only legitimate voters vote and that the general democratic principle of "one voter, one vote" is followed. As with benefits issuance, biometric systems can help in determining who is voting, that they are authorized to vote, and that they are not voting, or registering to vote, multiple times. Biometrics can also act as a fraud deterrent by, for example, facilitating the forensic identification of a person who attempts to vote using a deceased individual's credentials.

One example of biometrics in a voter registration application is the Bangladeshi Voter Registration Project. This project, run by the Bangladesh Army and Bangladesh Elections Commission, utilized fingerprint biometric technology to register voters for Bangladesh's 2008 general elections and to issue national identity cards. Four fingerprints were captured from each registrant and checked to see if they matched those captured from a prior registrant [7].

The Bangladeshi Voter Registration Project follows similar voter registration deployments conducted in countries like Mexico, Mozambique, and Nigeria. In Mexico, the Instituto Federal Electoral implemented a multi-biometric system that uses fingerprint and facial recognition to analyze historical voter rolls for duplicates, as well as to vet new voters against existing voter rolls [8].

Visa Shopping

In 2007, the European Parliament recognized a challenging problem facing several European Union member states: visa applicants rejected by a Schengen country were applying to other Schengen countries in the hopes of finding one that would issue them a visa. They were "visa shopping." In some cases, applicants would present forged documents as part of the visa application process.

To counter such attempts, the European Parliament established the Visa Information System (VIS). VIS is a database that contains fingerprint and face images and associates the collected biometrics with visa applicants' biographical data, as well as the dates and locations of application attempts [9]. Authorized officials responsible for border security can now better detect if a visa applicant has previously been rejected by another Schengen nation and is presenting falsified documents.

One result of the implementation of VIS has been that some visa applicants who have been rejected previously, or who have reason to believe they may be rejected, have mutilated their own fingers to avoid being processed against VIS. Others have attempted to perpetuate fraud by trying to alter their fingerprints using often painful processes with low chances of success, given that fingerprints extend beyond the epidermis.

Border Control

As with visa shopping deployments, border control deployments seek to protect national borders by determining with greater certainty who is entering (and, sometimes, exiting) a nation. These biometric applications help deter and expose entrance document fraud and identity fraud. They differ from visa shopping deployments insofar as they are oriented more towards identity and credential verification at the arrival and departure stages, rather than at the registration or application stages.

One example of a border control fraud reduction effort using biometrics is the United Arab Emirates' Iris Expellee Tracking System, deployed in 2003 and run by the Abu Dhabi Police. United Arab Emirates (UAE) officials were concerned about foreigners expelled from the UAE subsequently attempting to reenter the country after changing their name and/or nationality and then obtaining a new passport. The Iris Expellee Tracking System involves collecting iris images from all expelled foreigners. Arriving Passengers have their irises scanned at the UAE borders to verify that they were not formerly expelled. In 2005, the UAE reported catching approximately 32,850 previously expelled individuals [10].

Another program, the Canadian Passenger Accelerated Service System (CANPASS), uses iris recognition to allow preapproved, low risk travelers to clear Canadian customs and immigration without having to present documentation to border officials. Applicants who are approved to participate in CANPASS have their irises enrolled into the system. They then present their irises at designated kiosks at participating border environments (e.g., airports) for quick passage into Canada. This system not only helps reduce the

chance of forged documents passing inspection due to human error, it also allows border control officers to focus more on persons of greater interest who are more likely to be potential fraudsters.

Related Entries

- ▶ Asset Protection
- ▶ Binding of Biometric and User Data
- ▶ Biometric Encryption
- ▶ Consumer Recognition
- ▶ Forgery Sign
- ▶ Fraud Reduction, Applications
- ▶ Liveness and Anti-Spoofing
- ▶ Spoofing
- ▶ Time and Attendance

References

1. UK Department for Work and Pensions, Information Directorate: Fraud and Error in the Benefit System: October 2006 to September 2007 (May 2008)
2. Urbina, I.: Voter ID Battle Shifts to Proof of Citizenship, New York Times. http://www.nytimes.com/2008/05/12/us/politics/12vote.html?ref = opinion (12 May 2008). Accessed 2 Sept 2008
3. Samples, J.: The Motor Voter Act and Voter Fraud, CATO Institute. http://www.cato.org/testimony/ct-js031401.html (14 May 2001). Accessed 2 Sept 2008
4. Palm Print Recognition, National Science and Technology Council. http://www.biometrics.gov/Documents/PalmPrintRec.pdf. Accessed 2 Sept 2008
5. Front End Medicaid Fraud Reduction Pilot Program Based on Biometric Front-End System, Texas Health and Human Services Commission. http://www.hhsc.state.tx.us/OIE/RFP/FrontEnd/FingerImaging_RFI.pdf (30 June 2003). Accessed 2 Sept 2008
6. LGE Iris Tech Win in India Redefines Biometric Scalability. Findbiometrics.com. http://www.findbiometrics.com/article/115. Accessed 2 Sept 2008
7. BIO-key and Tiger IT Bangladesh Voter Registration Project Nearing Completion, BIO-key International, Inc. http://www.reuters.com/article/pressRelease/idUS121192 + 17-Jun-2008 + PRN20080617 (17 June 2008). Accessed 3 Sept 2008
8. The Mexican Instituto Federal Electoral (IFE): Mexico deploys multi-biometric voting system. Biometric Technol. Today. **14**(5), 3–4 (2006)
9. Kablenet: EU aims to stop 'visa shopping', The Register. http://www.theregister.co.uk/2007/06/08/schengen_visa_data/ (8 June 2007). Accessed 3 Sept 2008
10. Lieutenant Mohammad Al-Mualla: The UAE Iris Expellees Tracking and Border Control System. http://www.biometrics.org/bc2005/Presentations/Conference/2%20Tuesday%20September%2020/Tue_Ballroom%20B/Lt.%20Mohammad%20UAE2005.pdf. Accessed 4 Sept 2008
11. Study Report on Biometrics and E-Authentication. INCITS/M1 Ad Hoc Group on Biometrics in E-Authentication (2006)
12. Anderson, K.: Consumer Fraud in the United States: Second Survey. Federal Trade Commission, Washington, DC (2007)
13. Lee, V.: Transcript of Authentication Technologies FTC Proof Positive Workshop Session. Federal Trade Commission, Washington, DC (2007)
14. Progress in Tackling Benefit Fraud: Thirty-first Report of Session 2007–2008, UK House of Commons, Committee of Public Accounts. Accessed 2 June 2008

Fraud Reduction, Overview

Victor Minchih Lee
International Biometric Group, New York, NY, USA

Synonyms

Biometric fraud reduction; Fraud deterrence; Fraud mitigation; Identity theft reduction

Definition

Fraud is conventionally defined as the deliberate perversion or withholding of veracity to induce another to surrender something of value. In the context of biometrics, the item of value is typically an identity or a privilege associated with an identity.

Fraud can assume a variety of forms ranging from phishing to scams to hacking. In the specific case of biometrics, fraud can also consist of spoofing, or the presentation of an artifact designed to imitate a legitimate biometric.

Fraud reduction in a biometric context entails both the use of biometric technology to deter, inhibit, and mitigate fraud, as well as efforts to counter the exploitation of biometric system vulnerabilities through illegitimate submissions.

Introduction

With the increased reliance of modern society on technology, fraudsters have developed new exploitative techniques to prey upon the unsuspecting and the vulnerable. Internet merchants offer wares at hard-to-resist discounts, but never deliver their patrons any of the purchased goods. Sophisticated counterfeiters create fake currency and checks that are frequently difficult to distinguish from their genuine counterparts. Identity thieves send realistic, yet illegitimate, emails designed to harvest passwords and identity information from the careless and the inexperienced.

The cost of fraudulent activity, which includes tangible sums lost and expenditures to recover stolen goods, identities, or privileges, can be enormous. In 2005, British insurer Norwich Union estimated the cost of fraud in UK to be around £16 billion, or roughly 1.4% of UK's gross economic output (http://news.bbc.co.uk/1/hi/business/4463132.stm). In 2006, this number soared to £40 billion according to conservative UK government estimates (http://www.timesonline.co.uk/tol/news/uk/article633540.ece). The same year, identity theft alone victimized over 8 million people in the US to the tune of more than US $49 billion, according to the California Office of Privacy Protection. Unfortunately, loss due to fraudulent activities is likely to increase as fraudsters expand their reach thanks to increased globalization and rapid development of technology.

Technology, however, can be used both to abet, as well as combat, fraud. Biometric technologies and systems, especially, enable the deterrence, inhibition, and mitigation of fraud. For the purpose of this discussion, biometrics are defined as technologies that perform automated measurement of human physiological or behavioral characteristics to determine or authenticate identity.

Biometrics, thus, revolve around the concept of identity. Identity, in turn, is often seen as a proxy for trustworthiness, whether through linkage of identity to a historical and/or transactional record, or through connection of identity to a privilege or right.

Trustworthiness is the first victim of any fraudulent act. Security, robustness, and confidence of identity are therefore critical. Biometrics, if employed judiciously and with appreciation for the limitations and vulnerabilities of the technology, can satisfy such crucial needs.

Biometrics Vis-à-Vis Alternative Authentication/Identification Technologies

Biometric Advantages in Fraud Reduction

Currently several non-biometric methods and technologies exist to help combat fraud by serving an ▶ authentication or ▶ identification function. Examples of such technologies include smart cards, tokens, fobs, passwords, and personal identification numbers (PINs). Generally, these technologies can be divided into two categories: those based on something one has and those based on something one knows.

Biometrics introduces a third, complementary aspect: authentication and identification technologies based on something one *is*. By focusing on the elements that are inherent to an individual, biometrics offer additional protections that are unavailable or weaker through more traditional authentication/identification technologies. These include:

- Convenience
- Accountability
- Security

Convenience

Biometrics can obviate the carrying of tokens or cards that can be lost, misplaced, or – more saliently – stolen, leading to fraudulent access or unauthorized transactions. Biometrics can also eliminate the need to remember passwords or PINs. Often, people select simplistic passwords that can be easily guessed or hacked because they fear that they will forget more complex passwords.

Losing or forgetting a biometric, however, is considerably more difficult, especially for biometrics primarily dependent on physiological, rather than behavioral characteristics. Whereas one can suffer brain damage and forget how to sign one's name, rendering signature recognition useless, misplacing or forgetting a finger that is integrally tied to the rest of one's body is harder to accomplish. While violent criminals can go to extremes to remove a finger from a target so as to gain access to a fingerprint-secured facility, it is markedly more challenging to trick a target into unwittingly giving up such a personal feature.

Accountability

Biometrics are excellent technologies when transferability is of concern. Instead of relying on force or compulsion, fraudsters achieve success through convincing, coercion, and deception designed to encourage victims to surrender, or provide access to, a privilege, right, or item of value. Sophisticated scams, for example, can lead victims consciously and willingly to hand over precious access cards or passwords to perpetrators of fraud.

Biometric characteristics, however, are distinct and very personal. Their transference from one individual to another can be, as mentioned earlier, extremely challenging. This can contribute markedly to accountability, in addition to deterring and inhibiting fraud. If it is difficult for a fraudster to trick an individual into giving up their biometric, then any action taken that can be linked to that biometric is likely to have been undertaken by the legitimate possessor of the biometric in question. This makes it difficult to believe excuses in which a misdeed was allegedly committed by another who fraudulently obtained one's biometric characteristics.

With more traditional authentication/identification technologies, however, transferability can translate into reduced accountability. One fraudster could borrow access cards or passwords that allow him to take advantage of services or privileges intended for another. There could even be complicity in this effort – something that would take a high degree of personal sacrifice if biometrics were involved.

Biometrics can also add an element of accountability by deterring and inhibiting fraudulent attempts at establishing or relying on multiple identities. In the past, for instance, certain fraudsters with notorious histories of cashing bad checks would assume several identities so as to avoid the stigma and troubles accompanying their negative transactional histories. The introduction of biometric technologies and systems, however, has helped identify and address problems of multiple registrations by linking personal, biometric characteristics, rather than just nominal identities, to transactional histories and other historical records. This has also aided in the combating of fraudulent acts including multiple civil ID registrations and visa shopping.

Additionally, in some case, biometrics deter and mitigate acts of fraud by encouraging or necessitating the leaving behind of distinct, personal characteristics. For example, some prospective culprits may think twice before acting if they are aware that their criminal and fraudulent activity could potentially result in the leaving behind of biometric markers, such as their latent fingerprints. Those who proceed anyway and ignore the concern of enrolling in a biometric system and leaving behind an image or template of a distinct, personal characteristic could possibly be identified later and tracked by the biometrics they previously presented, a potential advantage for law enforcement and means of mitigating the severity and impact of a fraudulent act (e.g., by catching fraudsters before they are able to take advantage of the captured item of value or privilege).

Security

As described above, biometrics offers security and anti-fraud advantages over more traditional authentication/identification technologies with respect to identity transference, establishment of multiple fake identities, and loss or forgetting of credentials. They can render certain fraudulent activities – like phishing – almost irrelevant.

Biometric characteristics also provide additional anti-fraud security benefits thanks to their inherent nature; compared with passwords/PINs and cards/fobs/tokens, biometric characteristic is generally more difficult to capture, steal, replicate, and fake. Cards, for instance, are often designed to be robust, yet flexible enough that, in case they are lost, a replacement can be relatively easily created. PINs can be sniffed out through tracking or hidden monitoring technologies. They can also be readily discovered, in several cases, through brute force and trial-and-error techniques. Replication of a compromised PIN is then no more complicated than re-entering the newly revealed PIN.

It can be challenging, however, to create a replica of a biometric characteristic that has sufficient enough fidelity to work with a targeted biometric system. Creating a plausible fake iris, for example, often requires more effort than just copying electronic data onto a new smart card or retyping a password (in which cases the artifact will be identical to the genuine sample). This is due in part to liveness detection, a security function that is built into several biometric systems. Liveness detection, a fraud countermeasure, deters or

inhibits the presentation of artifacts, called spoofs, as legitimate biometric characteristics. Examples of liveness detection include: measurement of finger perspiration over time, 2D Fourier spectrum analyses, and behavioral reactions to cues (e.g., blinking upon command).

In addition, several biometric systems rely on templates, rather than full images of biometric characteristics, for reasons that range from privacy to cost to efficiency of data management and processing. Attempting to regenerate or reverse-engineer a complete biometric image from a select template is a very challenging, if not, at times, outright impossible, task. Also, trying alternative, brute-force techniques to recreate a biometric characteristic could take extremely lengthy and impractical periods of time, given the vast number and variability of components that make up many biometric characteristics.

Furthermore, biometric systems can often be costly, expensive, and technologically complex. Spending large sum of money to obtain a biometric device for study and identification of vulnerabilities and penetration points may not be cost effective. Likewise, even those who have the resources and know-how to create fake biometric characteristics, however, may find the effort of doing so to be cost-inefficient, especially when the value of the item or privilege being protected is outweighed by the cost of fraudulently obtaining it.

In some of the Panasonic's US offices, for example, hand geometry biometric readers are employed for time and attendance functions vis-à-vis custodial staff. Though several special effects and novelty item firms have the ability to create fake hand models, the cost and effort entailed in obtaining a suitable spoof would probably exceed the financial return of an extra hour of pay.

Authentication/Identification Trifecta

While biometrics offer significant advantages over more traditional and conventional authentication/identification technologies, it is important to note that this does not mean that biometrics should be employed *in lieu of* these other technologies. When issues of fraud, as well as security and protection of identity, are at stake, it may be optimal to leverage all proven options, especially given the potentially high cost of fraud and the ease with which fraud can often be committed.

Also, as will be discussed in the following section, biometrics have their own inherent vulnerabilities. These potential weaknesses can sometimes be mitigated by adopting complementary technologies which can provide an extra – if not necessarily equally effective – layer of defense. Where fraud is involved, the need for security may outweigh convenience and cost; such scenarios encourage reliance on an authentication trifecta that consists of:

- Something you *have*
- Something you *know*
- Something you *are*

Biometric Vulnerabilities

While biometric technologies can prove to be relatively robust and effective tools in fraud reduction through deterrence, inhibition, and mitigation, biometric systems themselves are not immune to fraudulent and exploitative attacks. These attacks can be classified according to three overarching categories:

- Input level attacks
- Processing and transmission level attacks
- Backend and storage level attacks

Input Level Attacks

Input Level Attacks generally fall into one of three categories:

- Spoofing attacks
- Bypassing attacks
- Overloading attacks

Spoofing attacks consist of attempts to deceive biometric system sensors into accepting an artifact as a legitimate biometric sample, typically for false enrollment, verification, or identification purposes. Spoofing attacks are usually considered to be attempts at breaking into biometric systems that are predominantly physiological in their focus. Biometric systems that are predominantly behaviorally based revolve less around the creation of spoof items and more around careful observation and practiced imitation of legitimate behavior.

Bypassing attacks consist of attempts to circumvent biometric system processes by creating artificial

Examples of spoof types for five established biometric modalities include:

Fingerprint	Face	Iris	Hand geometry	Voice recognition
Prostheses	Prostheses	Prostheses	Prostheses	Audio playback recordings
Props/Models/Gag items	Masks/Disguises	Video playback recordings	Props/Models/Gag items	Audio composite recordings
Photograph imitations	Photograph imitations	Photograph imitations		
Residual prints		Imprinted contact lenses		
Latent prints				

failures during enrollment or recognition so as to skip the biometric system altogether. One example of a bypassing attack would be to alter the quality of a biometric characteristic in such a way that a biometric system has difficulty in acquiring that characteristic. This could, for example, entail artificially filing down fingerprints so that there is a failure to enroll. The risk, as a result, is that an individual could then possibly be excused from biometric system recognition requirements and permitted to use a less robust authentication/identification system.

Closely related to bypassing attacks are variants called overloading attacks. In an overloading attack, a fraudster attempts to defeat or circumvent a biometric system by damaging or overwhelming the biometric sensor(s). Overloading attacks can range from flashing strobe lights against an optical sensor to presenting artificial heat sources to near-infrared-based sensors to short circuiting of sensitive sensors using liquids. As with bypassing attacks, the goal of an overloading attack is to either reduce the robustness, precision, and accuracy of the targeted biometric system and/or to encourage the substitution of the biometric recognition method with a less robust authentication/identification process and mechanism.

Processing and Transmission Level Attacks

Processing and transmission level attacks generally fall into one of three categories:

- Hacking
- Skimming/Sniffing
- Hill-Climbing

Processing and transmission level attacks are, strictly speaking, lesser acts of deception and fraud and more direct, technically based invasions. However, the result of success in any such attack on a biometric system could enable future acts of fraud, so it is important to be aware of these potential vulnerabilities.

Hacking, as herein defined, consists of electronically based attempts to penetrate a biometric system by altering the operation and functionality of the system through non-physical modifications and subterfuge (often at the code or system communications levels). A hacker could change the enrollment or recognition algorithms of a biometric system, lowering thresholds to accommodate less robust performance and security checks. They could program the system to forward them the copies of legitimate samples or instruct the system to allow them special, otherwise unauthorized, access.

Skimming and sniffing refers to techniques by which data is captured – often surreptitiously – during communication or processing of the information. Skimming devices, for example, could be designed to read and copy biometric data being submitted on a smart card to a biometric system for comparison against a live sample. This data could then be illegally replicated. Sniffing could occur if monitoring programs are put in place to capture data packets being sent from the capture sensor to the backend for verification.

Hill-climbing attacks first consist of the presentation of a test biometric sample to a biometric algorithm for comparison against an enrolled sample. A match score is then obtained and studied so that a new test sample can be presented for re-comparison and the achievement of a higher match score. This process is re-iterated until the biometric system's threshold has been discovered and is penetrable.

Backend and Storage Level Attacks

Backend and storage level attacks generally fall into two categories:

- Infiltration
- Implantation

As with the process and transmission level, the backend and storage levels are susceptible to malicious hacking. Skilled hacker-fraudsters could alter the permission levels tied to specific images or templates stored in databases. They could infiltrate the backend and alter the way biometric data that is classified and stored. More of concern, they could perhaps steal biometric characteristics data and try to generate spoofs using the information captured.

In addition, acts of fraud can be facilitated if fraudsters are able to gain unauthorized or complicit access to backend and storage databases of biometric information to perform acts of implantation. In this attack, fraudsters might implant their own biometric characteristics into a targeted biometric system's database. By doing so, fraudsters would be able to appear as legitimately authorized individuals with free access to the rights or privileges otherwise secured by the biometric system.

Countermeasures

In order to counter – or at least inhibit – the three aforementioned types of attacks, certain countermeasures can be enacted. These countermeasures can be classified according to the level of attack they are best suited to address.

At the input level, spoofing is typically counteracted by the attempt to determine whether a live, real human sample is being presented to the capture device. This is, as mentioned earlier, called liveness detection and is based on the assumption that, with the exception of some cadaver recognition applications, a legitimate biometric will always be presented by the live possessor of that biometric characteristic.

As for bypassing and overloading attacks, countermeasures include increased ruggedization of capture devices and sensor equipment, conscientious form factor design (e.g., creating shielding from external light sources that could be potentially malevolent), supervision of enrollment and recognition submission processes, as well as rigorous fallback procedures and processes. After all, those who seek to accomplish fraud will often target the weakest link. If this means taxing a biometric system out so that, for example, access to a secure facility can be obtained through a potentially more fallible human guard inexperienced at identifying fake identity documents, which will often be the strategy of choice for motivated fraudsters.

At the processing and transmission levels, countermeasures may entail proven information systems and information technology security techniques, such as the use of firewalls and encryption. After all, at a certain level, biometric data is often converted into streams of digital data that should be accorded no less than the rudimentary security protections already commonplace for digital information that is less personally sensitive. In addition, best practices should be implemented, such as requiring the use of data transmission shields (that limit the range at which data can be sniffed from contactless smart chips), as well as strict limitation of access to matching score data.

At the backend and storage levels, highly advisable countermeasures include firewalls, as well as extensive auditing functions and logs of modifications executed (whether they are additions, subtractions, or alterations of biometric data). A best practice countermeasure would also be the frequent, though not necessarily habitual or scheduled, review of random images and templates for evidence of tampering, alteration, missing presence, or unexpected presence.

In order to further deter or inhibit the abuse of biometric systems by fraudsters, biometric systems may also be designed with the following four countermeasures:

- ► Multifactor or ► multimodal authentication requirements
- Randomization of modality
- System challenges
- Emphasis on internal/subcutaneous characteristics

By adopting multifactor or multimodal systems, deployers increase the challenge for fraudsters by requiring them to defeat several disparate systems for which the optimum exploitation and penetration techniques may be very different. Though there is a convenience tradeoff, the security that accrues can be significant, particularly when security of identity is at stake. The main caveat, however, is that potential fraudsters are not encouraged to pretend to be unable to use one of the modalities so as to simplify their, say, spoofing task.

Examples of liveness detection methods for five established biometric modalities include:

Fingerprint	Face	Iris	Hand geometry	Voice recognition
Spectroscopic analysis	Reactivity to Cues, Commands, and stimuli (e.g., – blinking)	Photonic and spectrographic analysis	Required contact with specifically- placed prongs	Recitation of randomly generated passphrases
Temporal variation in perspiration		Reactivity to stimuli (e.g., – pupil dilation)		
		Ink/Dye detection		
		Timestamping and byte scrambling		

To provide a little more balance between convenience and security, a multimodal system could still be employed, but only with one or two randomly selected biometric modalities required for authentication/identification. Variations could also be introduced within a single modality (e.g., requiring submission of a right index finger, one day, and submission of a left thumb on the next day).

The authentication/identification systems could also be designed to issue randomized as well as cued challenges – even if the original submission would otherwise have been acceptable. At the very least, this implementation would provide the opportunity to obtain two biometric samples. Where the samples are unusually similar, extra caution might be merited in case a spoof is involved, as the likelihood that a person will be as infallible as to place their biometric so consistently is slim.

Finally, biometric systems can be deployed and designed so as to focus on internal or subcutaneous characteristics that are generally much more difficult to capture surreptitiously, as well as to forge or modify.

Biometrics and Fraud: Looking to the Future

While a lot of focus has been placed on the design and utilization of biometrics to deter, inhibit, and – to a lesser degree – mitigate non-biometric fraud, comparatively little attention has been paid to the consequences and implications when fraudsters are successful in compromising biometric data. Because biometric characteristics are so intrinsic and relatively immutable, this is an issue of particular concern that can impact the successful deployment of the technologies.

Cancelable/Changeable Biometrics

To address concern over the immutability of biometrics, research has been conducted by entities like IBM and the Korean Biometrics Engineering Research Center into cancellable biometrics, also known as changeable biometrics. The high-level concept of such research has been to look into altering biometric data that is captured before it is actually fully processed and stored in template form. In this way, a compromised biometric can theoretically be revoked and a new algorithm can generate a novel distortion of the affected individual's biometric characteristic – essentially giving them a new biometric.

However, one should keep in mind that if a fraudster is able to get hold of the original source biometric characteristic (or an equivalent spoof), this approach would not suffice, as the fraudster would then still be able to regenerate a new cancelable/changeable biometric characteristic just as easily as the legitimate bearer of the original source biometric.

Nominal Identities Versus Biometric Identities

As biometric systems increasingly protect sensitive data and items or access privileges of high value, the incentive fraudulent activity to exploit them will also increase. And at some point, as the case has been with virtually every major security technology in the past, biometric data will be compromised.

One of the important ways in which the impact of such compromise can be mitigated is to sever, whenever feasible and reasonable, the connection between

an individual's nominal identity and their biometric identity. If, for instance, a deployment merely requires a determination as to whether a given individual, represented by their biometric characteristics, should be granted access to a given secure location, then there is no need to link permanently the individual's name and background information to their biometric data after an initial background check has been conducted.

Whereas names have often served as proxies for trustworthiness or transactional histories, biometrics can now serve this purpose going forward. With biometrics there is also the possibility of selecting different biometric aspects for accreditation or validation given each distinct application or deployment. A single biometric characteristic, thus, has the flexibility to serve in a variety of functions that process that biometric differently – without making that biometric into a universal identifier rife with the problems of overuse that have plagued the US social security number.

In the scenario described above, if a fraudster compromises one biometric system, the damage is mitigated insofar as other systems and deployments may still be protected, in addition to sensitive and private information tied to one's nominal identity.

Valuing Biometric Data

One of the remaining challenges with respect to biometrics and fraud is the determination of how to value biometric data. This is especially important as fraudsters increasingly target not just data protected by biometrics, but biometric data, itself.

Traditionally, items have been valued based on three factors:

- Scarcity
- Uniqueness
- Demand

With biometrics, however, such a framework for assessing value is of little use: virtually each and every given biometric characteristic is inherently distinct (if not unique), scarce, and of high demand for both the possessor and potential imposters/fraudsters. It would seem, therefore, that all biometric characteristics should be deemed priceless or at least assigned extremely lofty values.

However, this would be impractical in an age of risk calculations and need by insurance companies, governments, and other entities realistically to quantify the impact and cost of fraud. Therefore, valuation of a biometric is best conducted according to a different set of three factors:

- Value of the Biometrically-Protected Item or Privilege
- Range of Utility
- Spoofability

In addition, whenever a biometric system is designed, careful consideration needs to be taken as to whether templates or images should be used. Generally, images will be more valuable from perspectives concerned with forensics, interoperability, and scalability. Templates, however, will be more desirable from an identity-protecting perspective. Thus, from a fraud reduction perspective, the guiding principle should be that templates, which are more limited than images, normally, should be employed whenever possible in lieu of images. To achieve this balance, a negative incentive should be implemented such that there will be stiffer legal penalties for compromised biometric image data versus biometric template data.

Related Entries

- ▶ Asset Protection
- ▶ Binding of Biometric and User Data
- ▶ Biometric Encryption
- ▶ Consumer Recognition
- ▶ Forgery Sign
- ▶ Fraud Reduction, Applications
- ▶ Liveness and Anti-Spoofing
- ▶ Spoofing
- ▶ Time and Attendance

References

1. Study Report on Biometrics and E-Authentication. INCITS/M1 Ad Hoc Group on Biometrics in E-Authentication (2006)
2. Anderson, K.: Consumer Fraud in the United States: Second Survey. Federal Trade Commission, Washington, DC (2007)
3. Lee, V. et al.:Transcript of Authentication Technologies FTC Proof Positive Workshop Session. Federal Trade Commission, Washington, DC (2007)

4. Fraud and Error in the Benefit System: October 2006 to September 2007. UK Department for Work and Pensions, Information Directorate (May 2008)
5. Progress in Tackling Benefit Fraud: Thirty-first Report of Session 2007–2008. UK House of Commons, Committee of Public Accounts (2 June 2008)
6. Counting the cost of UK fraud, BBC News, 24 November 2005. http://news.bbc.co.uk/1/hi/business/4463132.stm (29 September 2008)
7. Woolcock, N.: Cost of fraud spirals to £40bn, TimesOnline, 9 September 2006. http://www.timesonline.co.uk/tol/news/uk/article633540.ece (29 September 2008)
8. Ratha, N. et al.: Cancelable Biometrics: A Case Study in Fingerprints (2006)
9. Unisys. Consumers Worldwide Overwhelmingly Support Biometrics for Identity Verification, 26 April 2006. Press release (30 September 2008)
10. Use of Biometric Identification Technology to Reduce Fraud in the Food Stamp Program, United States Department of Agriculture, Food and Nutrition Service, December 1999. http://www.fns.usda.gov/oane/MENU/Published/fsp/FILES/ProgramIntegrity/biomeval.htm (30 September 2008)
11. Identity Fraud: Prevalence and Links to Alien Illegal Activities, United State General Accounting Office, 25 June 2002. http://www.gao.gov/new.items/d02830t.pdf (30 September 2008)
12. Martinez-Diaz, M. et al.: Hill-Climbing and Brute-Force Attacks on Biometric Systems: A Case Study in Match-on-Card Fingerprint Verification, http://fierrez.ii.uam.es/docs/2006_ICCST_HillClimbingAttackMoC_Martinez.pdf (30 September 2008)
13. Schuckers, S. et al.: Issues for Liveness Detection in Biometrics. http://www.biometrics.org/html/bc2002_sept_program/2_bc0130_DerakhshabiBrief.pdf (30 September 2008)

Freeman Chain Code (FCC)

Freeman Chain Code (FCC) is a compact method for representing the contours of an object, first made popular by Herbert Freeman.

► Hand Data Interchange Format, Standardization

Function Creep

This refers to the use of data beyond the purposes originally intended at the time of data collection. For biometrics, this usually means using the data for purposs other than identification, as when a face image is used to determine gender or ethnicity.

► Privacy Issues

Fundamental Frequency, Pitch, F0

The fundamental frequency or F0 is the frequency at which vocal chords vibrate in voiced sounds. This frequency can be identified in the sound produced, which presents quasi-periodicity, the pitch period being the fundamental period of the signal (the inverse of the fundamental frequency). Pitch is more often used to refer to how the fundamental frequency is perceived.

► Speech Analysis

Fusion Network Topology

The network architecture including sensors, communication channels, and fusion processing. The fusion processing may be distributed due to physical constraints in the system. If a communication channel between two sensors is long, it may be beneficial to fuse all the sensors at one end of the channel so that only the single fused decision is sent through the channel. The topology is directly impacted by the physical layout of the sensor network.

► Fusion, Decision Level

Fusion, Biometric

See Multi-biometrics.

► Biometric Algorithms
► Multibiometrics and Data Fusion Standardization
► Multiple Experts

Fusion, Confidence Level

► Fusion, Score-Level

Fusion, Data Level

► Fusion, Sensor-Level

Fusion, Decision-Level

Lisa Osadciw, Kalyan Veeramachaneni
Syracuse University, Syracuse, NY, USA

Synonyms

Distributed detection; Distributed inference making; Multiple classifier fusion; Statistical signal processing

Definition

Decision level fusion falls under a broader area known as distributed detection systems and is the process of selecting one hypothesis from multiple M hypotheses given the decisions of multiple N sensors in the presence of noise and interference. In biometrics, decision level fusion creates a single decision from typically two hypotheses, imposter or genuine user, from multiple biometric sensor decisions, which may or may not be identical sensors. Often, decision level fusion is implemented to save communication bandwidth as well as improve decision accuracy. A statistical performance model for each biometric sensor is needed a priori to support the system wide optimization in terms of two error rates: false acceptance rate, admitting an imposter, and false rejection rate, rejecting the genuine user. A weighted sum of these two errors is a useful objective function. This provides the designer with the flexibility to weigh one error more than the other error. Decision level fusion may be done at one processor, centrally, or at multiple processors, distributed.

Introduction

In biometric decision level fusion, the biometric sensors send their final decisions through a communication network that finally fuses these decisions at a fusion center. Optimal decision fusion theory can be applied to these problems. In distributed detection systems, the number of decisions a sensor can make varies as well as the ► fusion network topology. It may be more advantageous to fuse a few sensors at a local node before transmitting the information over a long distance to the final fusion processor [1]. This complicates the fusion problem by introducing different fusion network topologies. Decision level fusion remains at the foundation of the problem, however.

The decision level fusion problem in the biometric area is typically one in identifying the user as a genuine user or an imposter with the final decision made by a central fusion processor [2–5]. This is referred to as a ► parallel fusion network. The advantages of fusion are twofold. The first advantage is a more accurate final decision by using multimodal, multiple and diverse, biometric sensors, which provide significantly more information to base a decision. Secondly, communication bandwidth needs, which are great as more sensors are networked, remain relatively constant if the decisions instead of the full observation or measurement are communicated.

The fusion accuracy of the sensor decisions relies on the accuracy of the statistical models for the sensors and an optimally designed fusion rule. The biometric verification problem may be posed as a ► binary hypothesis testing problem with the match score(s) serving as observations. The two hypotheses are

H_0: Imposter Identified and
H_1: Genuine User Identified

Probability of false alarm,

$$P_{\mathrm{FA}} = P(u = 1|H_0) \tag{1}$$

and probability of false rejection,

$$P_{\mathrm{FR}} = P(u = 0|H_1) \tag{2}$$

In the Bayesian formulation, these two errors are weighted by costs and summed into a single cost function called the Bayesian risk function. The Bayesian risk function is

$$R = P(H_0) \times C_{\mathrm{FA}} \times P_{\mathrm{FA}} + P(H_1) \times C_{\mathrm{FR}} \times P_{\mathrm{FR}}, \tag{3}$$

where $P(H_0)$, a priori probability of an imposters, $P(H_1)$ a priori probability of a genuine user C_{FA}, cost of false acceptance, and C_{FR}, cost of false rejection. In the worst-case scenario, one assumes equal a priori probabilities. Thus, we get

$$R = C_{FA} \times P_{FA} + C_{FR} \times P_{FR}. \quad (4)$$

We can rewrite Eq. 4, using a single cost factor by replacing the cost of false acceptance by

$$C_{FA} = 2 - C_{FR}. \quad (5)$$

This simplifies the problem to one design parameter to optimize if the a priori probabilities of genuine users and imposters are assumed to be fixed.

Decision Level Fusion with Single Bit Information

Often Gaussian distribution functions are used as the statistical sensor models for a binary hypothesis problem. Each sensor has a different Gaussian distribution function as shown in Fig. 1 for each hypothesis: genuine user and imposter. Higher observation values are typically associated with a positive user identification or the genuine user hypothesis. This leads to the distribution on the right side of the plot in Fig. 1. The imposter has a lower mean. Sensor 1 must measure a score that exceeds a threshold for comparison purpose to decide if it has the genuine user. The error rates are simply the areas under the distribution corresponding to the opposite hypothesis or wrong side of the threshold. False acceptance probability is the area to the right of the threshold under the imposter distribution. False rejection probability is the area to the left of the threshold under the genuine user distribution. A single bit of 1 denotes that the user is detected while the 0 is for the imposter [7].

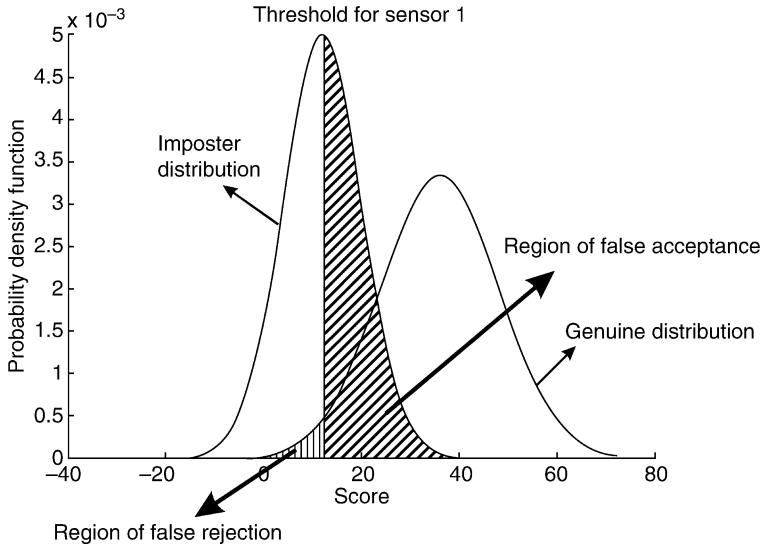
The threshold divides the entire decision region into region of acceptance and region of rejection. If a user's matching score happens to fall above the thresholds, he/she is considered as genuine. If the user's matching score falls below threshold he/she is considered as imposter [1, 2, 5, 6]. This process using the threshold, λ_i, for sensor i is given by

$$u_i = \begin{cases} 1, & y_i >= \lambda_i \\ 0, & y_i < \lambda_i \end{cases} \quad \forall i \quad (6)$$

Let $[u] = [u_1, u_2, \ldots, u_n]$, be the combined vector of decisions represented by 1s and 0s for all the sensors. These decisions are combined using a fusion rule of

$$u_f = f([u]). \quad (7)$$

The complete set of fusion rules for the 2-sensor case is given in Table 1 [1]. There are 16 possible rules for 2 sensors or $(2^2)^N$ with N sensors as 2. The fusion rule can be written as a 4-bit vector, where each bit represents the final fused decision given the



Fusion, Decision-Level. Figure 1 Illustration of thresholding process, Decision Regions, and Error Regions, for given Gaussian conditional density functions.

Fusion, Decision-Level. Table 1 All Possible Fusion Rules for 2 Sensors

$u_1\ u_2$	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}
0 0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0 1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1 0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1 1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

sensor decisions, $[u_1\ u_2]$. Since, there are two hypotheses, the fusion rules are based on Boolean logic. For example, in Table 1, f_2 is a rule based on AND logic. The final decision is "1" only when both the sensors say "1" and is "0" otherwise. Similarly, f_9 is "NAND", and f_8 is "OR".

For 3 sensors, there are 8 possible vectors requiring a fused decision. Hence, the fusion rule is 8 bits long, and the number of possible fusion rules for this problem is $(2^2)^3$ or 64 rules for 3 sensors.

The error probabilities as in Eqs. 1 and 2 for the entire system after fusion is estimated using

$$GP_{\mathrm{FA}} = \sum_{[u]} P(u_f = 1|[u], H_0) P([u]|H_0). \tag{8}$$

Assuming independence, Eq. 8 can be calculated using the statistical models and

$$\mathrm{GP}_{\mathrm{FA}} = \sum_{[u]} P(u_f = 1|[u], H_0) \prod_{i=1}^{n} P(u_i|H_0)$$

$$= \sum_{[u]} P(u_f = 1|[u], H_1) \prod_{i=1}^{n} \int_{y_i} P(y_i|H_0) dy_i \tag{9}$$

In case of correlation [8], however, the product disappears resulting in a multivariate integral or

$$= \sum_{[u]} P(u_f = 1|[u], H_0) \left(\int_{y_1} \int_{y_2} \cdots\cdots\cdots \int_{y_n} f_{Y_1, Y_2, \cdots, Y_n} (y_1, y_2, \ldots, y_n|H_0)\, dy_1 dy_2, \ldots, dy_n \right) \tag{10}$$

Similarly,

$$\mathrm{GP}_{\mathrm{FR}} = \sum_{[u]} P(u_f = 0|[u], H_1) P([u]|H_1). \tag{11}$$

Assuming independence, (11) can be calculated using

$$\mathrm{GP}_{\mathrm{FR}} = \sum_{[u]} P(u_f = 0|[u], H_1) \prod_{i=1}^{n} P(u_i|H_1). \tag{12}$$

In case of correlation, the multivariate integral arises as before giving

$$= \sum_{[u]} P(u_f = 0|[u], H_1) \left(\int_{y_1} \int_{y_2} \cdots\cdots\cdots \int_{y_n} f_{Y_1, Y_2, \ldots, Y_n} (y_1, y_2, \ldots, y_n|H_1)\, dy_1 dy_2, \ldots, dy_n \right) \tag{13}$$

This multivariate integral can only be calculated using numerical methods. Since there are 2^N combinations of local decisions for N sensors, this integral must be evaluated $2^N - 1$ times to estimate each error. This operation can be very expensive computationally as the number of sensors increases. An alternative is using the Bahadur–Lazarfeld expansion, which enables the estimation of the error probabilities using "$n-1$" evaluations of integrals [8].

The Bayesian risk function is now given by,

$$R = P_0 C_{10} P(u_0 = 1|[u], H_0) + P_1 C_{01} P(u_0 = 0|[u], H_1). \tag{15}$$

Optimal Fusion Rule

For independent sensors, however, the optimal fusion rule is the ▶ likelihood ratio test [5]. For fixed thresholds, the optimal fusion rule can be obtained by using the likelihood ratio as in

$$\frac{P(u_1, u_2, u_3 \cdots\cdots\cdots u_n|H_1)}{P(u_1, u_2, u_3 \cdots\cdots\cdots u_n|H_0)} \underset{u_0 = 0}{\overset{u_0 = 1}{\gtrless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}. \tag{16}$$

Fusion, Decision Level. Table 2 Means and Standard Deviations of the Gaussian Distributions Under both the Hypothesis for Different Sensors

Hypothesis/Paremeter	H_0/μ_0	H_0/σ_0	H_1/μ_1	H_1/σ_1
Sensor 1	47.375	43.864	144.514	12.843
Sensor 2	67.755	52.633	251.209	23.008

Fusion, Decision-Level. Table 3 Thresholds for the 3 Sensors for Single Bit Information

Sensor	Threshold	FAR	FRR
1	95.945029260756	0.13409060569213	0.00007790554000
2	185.799498919171	0.01245661704014	0.00223574300430

$P(u_1, u_2, u_3 \cdots\cdots\cdots u_n|H_h)$ in (16) can be replaced by $\prod_{i=1}^{n} \int_{y_i} P(y_i|H_h)dy_i$ in case of independence or $\int_{y_1}\int_{y_2}\cdots\cdots\cdots\int_{y_n} f_{Y_1,Y_2,\ldots,Y_n}(y_1, y_2, \ldots, y_n|H_h)dy_1\,dy_2, \ldots, dy_n$ in case of correlation. Optimal fusion rule can be employed when the thresholds are fixed. The optimal fusion rule as in Eq. 16 minimizes the Bayesian risk function.

In the case of independence [5], the optimum rule simplifies to

$$\sum_{i=1}^{N}\left[u_i \log\left(\frac{1-F_{FR}}{F_{AR}}\right) + (1-u_i)\log\left(\frac{F_{\mathrm{FR}}}{1-F_{\mathrm{AR}}}\right)\right] \begin{array}{c} u_f = 1 \\ > \\ < \\ u_f = 0 \end{array} \log\left(\frac{C_{\mathrm{FA}}}{2-C_{\mathrm{FA}}}\right). \quad (17)$$

Independent Pair of Biometric Sensors

Consider two sensors with conditional distributions under both the hypotheses given by the familiar Gaussian distribution. A Gaussian distribution is characterized by $N(\mu, \sigma)$ with a different mean, μ, and standard deviation, σ, for each hypothesis as mentioned earlier. Table 2 gives the parameters of the Gaussian distributions used for the 2 sensors in this example. In Table 3, the threshold that achieves the false alarm rate and false rejection rate given is specified for both sensors. Using these thresholds as well as the error rates in the optimal fusion rule of Eq. 17, we give the rules in the right column for the specified costs in the left. Thus, different rules become optimum as the error rates are weighted differently. If the sensor is replaced with a more accurate biometric sensor, the rule selection will change. Finally, if the sensors are correlated, the original rule in Eq. 16 must be applied and performance computed accordingly.

Fusion, Decision-Level. Table 4 Optimal Fusion Rule Under Assumption of Independence

CFA	Optimal Fusion Rule
0.2	Majority Voting Rule
0.6	Majority Voting Rule
1	(1 OR 2) AND 3
1.2	(1 OR 2) AND 3
1.5	(1 OR 2) AND 3
1.9	(1 OR 2) AND 3

Related Entries

▶ Multiple Experts

References

1. Varshney, P.K.: Distributed Detection and Data Fusion, Springer. Springer-Verlag, New York, Inc (1997)

2. Prabhakar, S., Jain, A.: "Decision-level Fusion in Fingerprint Verification", Pattern Recognit. **35**, 861–874 (2002)
3. Osadciw, L., Varshney, P., Veeramachaneni, K.: "Optimum Fusion Rules for Multimodal Biometric Systems", Foresti, G.L., Regazzoni, C.S., Varshney, P.K. Chap. 15, Multisensor Surveillance Systems: The Fusion Perspective, Kluwer (2003)
4. Veeramachaneni, K.: "An Evolutionary Algorithm Based Dynamic Thresholding for Multimodal Biometrics" Masters thesis, School of Electrical and Computer Engineering, Syracuse University (2003)
5. Chair, Z., Varshney, P.K.: "Optimal Data Fusion in Multiple Sensor Detection Systems", IEEE Trans. Aerosp. Electron. Syst. **22**(1), 98–101 (1986)
6. Tang, Z.-B., Pattipati, K.R., Kleinman, D.L.: "An Algorithm for Determining the Decision Thresholds in a Distributed Detection Problem". IEEE Trans. Syst. Man Cybern. **21**, 231–237 (1991)
7. Veeramachaneni, K., Osadciw, L., Varshney, P.: «Adaptive Multimodal Biometric Management Algorithm «, IEEE Trans. Syst. Man Cybern. **35** (2005)
8. Kam, M., Zhu, Q., Gray, W.S.: "Optimal Data Fusion of Correlated Local Decisions in Multiple Sensor Detection Systems," IEEE Trans. Aerosp. Electron. Syst. **28**, 916–920 (1992)

Fusion, Feature-Level

Arun Ross
West Virginia University, Morgantown, WV, USA

Synonym

Feature Fusion

Definition

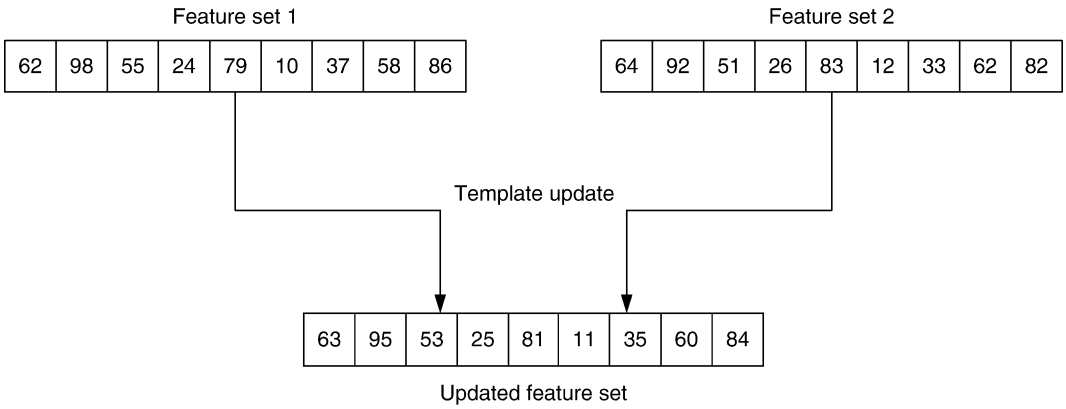
In feature-level fusion, the feature sets originating from multiple biometric sources are consolidated into a single feature set by the application of appropriate feature normalization, transformation, and reduction schemes. The primary benefit of feature-level fusion is the detection of correlated feature values generated by different biometric algorithms thereby identifying a compact set of salient features that can improve recognition accuracy. Eliciting this feature set typically requires the use of ▶ dimensionality reduction methods and, therefore, feature-level fusion assumes the availability of a large number of training data. Feature-level fusion algorithms can also be used for template update or template improvement.

Introduction

Feature level fusion is an example of an early fusion strategy, i.e., the biometric evidence from multiple sources are consolidated *before* invoking the matcher. In this scheme, multiple feature sets are integrated in order to generate a single template that is expected to be more robust than the individual feature sets. When the feature sets to be integrated are homogeneous (e.g., multiple measurements of a person's hand geometry), a single feature vector can be computed as a weighted average of the individual feature sets. When the feature sets are nonhomogeneous (e.g., features of different biometric modalities like face and hand geometry), they can be concatenated to form a single feature set. Feature selection schemes are employed to reduce the dimensionality of the ensuing feature set [1]. Concatenation is not possible when the feature sets are incompatible (e.g., fingerprint minutiae and eigenface coefficients).

If the feature sets to be combined originate from the same feature extraction algorithm (thus, a single modality is assumed) then feature level fusion can be used for template update or template improvement as discussed in the following section.

1. *Template update*: The template in the database can be updated based on the evidence presented by the current feature set in order to reflect (possibly) permanent changes in a person's biometric. Hand geometry systems use this process to update the geometric measurements stored in the database in order to account for changes in an individual's hand over a period of time. A simple scheme would be to take the average of the two feature vectors corresponding to the two instances of the biometric signal and use the average feature vector as the new template (Fig. 1).
2. *Template improvement*: In the case of fingerprints, the minutiae information available in two impressions can be combined by appropriately aligning the two prints and removing duplicate minutia thereby generating a larger minutia set. This process, known as template improvement, can also be used to remove spurious minutiae points that may be present in a feature set. While template update is used to accommodate temporal changes in a person's biometric, the purpose of template improvement is to increase the number of features (*and* decrease the number of spurious features) in the template.



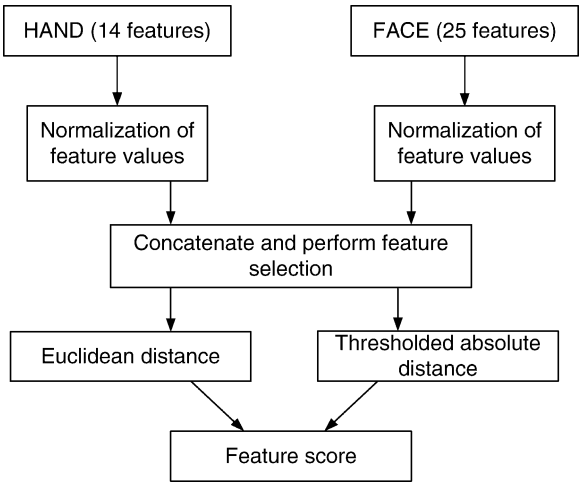
Fusion, Feature-Level. **Figure 1** A template update procedure may be viewed as a feature fusion scheme. In this example, the nine-dimensional feature set of a user ("Feature Set 1") is updated based on the evidence presented by the current feature set ("Feature Set 2"), via the averaging scheme.

Several template improvement algorithms have been discussed in the literature for fingerprints. Jiang and Ser [2] propose a template improvement scheme where a reliability measure is associated with each extracted minutia point. This reliability measure is updated as minutiae evidence from newly acquired impressions is made available. The parameters of a minutia point (i.e., its x-y location and orientation) are updated via a weighted average scheme; even the "type" of the minutiae (i.e., ridge-ending or ridge-bifurcation) is altered if necessary. Template improvement is applicable only when the new fingerprint impression is accurately aligned with the stored one. The authors use the match score to determine if two impressions are accurately aligned. During the verification stage, only those minutia points whose reliability measure is above a certain threshold are used in the matching process. The authors show that their scheme results in (1) the elimination of spurious minutiae points, (2) the addition of missed minutiae points, (3) the relabeling of incorrect minutiae types and, consequently, (4) a general improvement in matching performance. Other algorithms for minutiae template improvement have been discussed in [3, 4].

Feature Fusion Scheme

How does one consolidate feature sets originating from different algorithms and modalities? Feature level fusion is difficult to achieve in such cases because of the following reasons:

1. The relationship between the feature spaces of different biometric systems may not be known.
2. The feature sets of multiple modalities may be incompatible. For example, the minutiae set of fingerprints and the eigen-coefficients of face are irreconcilable. One is a variable length feature set (i.e., it varies across images) whose individual values parameterize a minutia point; the other is a fixed length feature set (i.e., all images are represented by a fixed number of eigen-coefficients) whose individual values are scalar entities.
3. If the two feature sets are fixed length feature vectors, then one could consider augmenting them to generate a new feature set. However, concatenating two feature vectors might lead to the ▶ curse-of-dimensionality problem [5] where increasing the number of features might actually degrade the system performance especially in the presence of small number of training samples. Although the curse-of-dimensionality is a well known problem in pattern recognition, it is particularly pronounced in biometric applications because of the time, effort and cost required to collect large amounts of biometric (training) data.
4. Most commercial biometric systems do not provide access to the feature sets used in their products. Hence, very few biometric researchers have



Fusion, Feature-Level. Figure 2 The procedure adopted by Ross and Govindarajan [1] to perform feature level fusion.

focused on integration at the feature level and most of them generally prefer fusion schemes that use match scores or decision labels.

If the length of each of the two feature vectors to be consolidated is fixed across all users, then a feature concatenation scheme followed by a dimensionality reduction procedure may be adopted. Let $\boldsymbol{X} = \{x_1, x_2, \ldots, x_m\}$ and $\boldsymbol{Y} = \{y_1, y_2, \ldots, y_n\}$ denote two feature vectors ($\boldsymbol{X} \in \boldsymbol{R}^m$ and $\boldsymbol{Y} \in \boldsymbol{R}^n$) representing the information extracted from two different biometric sources. The objective is to fuse these two feature sets in order to yield a new feature vector, $\boldsymbol{Z}$, that would better represent an individual. The vector $\boldsymbol{Z}$ of dimensionality k, $k < (m + n)$, can be generated by first augmenting vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, and then performing feature selection or feature transformation on the resultant feature vector in order to reduce its dimensionality. The key stages of such an approach are described as follows (also see Fig. 2).

Feature Normalization

The individual feature values of vectors $\boldsymbol{X} = \{x_1, x_2, \ldots, x_m\}$ and $\boldsymbol{Y} = \{y_1, y_2, \ldots, y_n\}$ may exhibit significant differences in their range as well as form (i.e., distribution). Augmenting such diverse feature values will not be appropriate in many cases. For example, if the x_i's are in the range $[0, 100]$ while the y_i's are in the range $[0, 1]$, then the distance between two augmented feature vectors will be more sensitive to the x_i's than the y_i's. The goal of feature normalization is to modify the location (mean) and scale (variance) of the features values via a transformation function in order to map them into a common domain. Adopting an appropriate normalization scheme also helps address the problem of outliers in feature values. While a variety of normalization schemes can be used, two simple schemes are discussed here: the min–max and median normalization schemes.

Let x and x' denote a feature value before and after normalization, respectively. The min–max technique computes x' as

$$x' = \frac{x - \min(F_x)}{\max(F_x) - \min(F_x)}, \tag{1}$$

where F_x is the function which generates x, and $min(F_x)$ and $max(F_x)$ represent the minimum and maximum of all possible x values that will be observed, respectively. The min–max technique is effective when the minimum and the maximum values of the component feature values are known beforehand. In cases where such information is not available, an estimate of these parameters has to be obtained from the available set of training data. The estimate may be affected by the presence of outliers in the training data and this makes min–max normalization sensitive to outliers. The median normalization scheme, on the other hand, is relatively robust to the presence of noise in the training data. In this case, x' is computed as

$$x' = \frac{x - median(F_x)}{median(|(x - median(F_x))|)}. \tag{2}$$

The denominator is known as the Median Absolute Deviation (MAD) and is an estimate of the scale parameter of the feature value. Although, this normalization scheme is relatively insensitive to outliers, it has a low efficiency compared to the mean and standard deviation estimators. Normalizing the feature values via any of these techniques results in modified feature vectors $\boldsymbol{X}' = \{x_1', x_2', \ldots x_m'\}$ and $\boldsymbol{Y}' = \{y_1', y_2', \ldots y_m'\}$. Feature normalization may not be necessary in cases where the feature values pertaining to multiple sources are already comparable.

Feature Selection or Transformation

Augmenting the two feature vectors, $\boldsymbol{X}'$ and $\boldsymbol{Y}'$, results in a new feature vector, $\boldsymbol{Z}' = \{x'_1, x'_2, \ldots x'_m, y'_1, y'_2, \ldots y'_n\}$, $\boldsymbol{Z}' \in \boldsymbol{R}^{m+n}$. The curse-of-dimensionality dictates that the augmented vector of dimensionality $(m + n)$ need not necessarily result in an improved matching performance compared to that obtained by $\boldsymbol{X}'$ and $\boldsymbol{Y}'$ alone. The feature selection process is a dimensionality reduction scheme that entails choosing a minimal feature set of size k, $k < (m + n)$, such that a criterion (objective) function applied to the training set of feature vectors is optimized. There are several feature selection algorithms in the literature, and any one of these could be used to reduce the dimensionality of the feature set $\boldsymbol{Z}'$. Examples include sequential forward selection (SFS), sequential backward selection (SBS), sequential forward floating search (SFFS), sequential backward floating search (SBFS), "plus l take away r" and ▶ branch-and-bound search (see [6, 7] for details). Feature selection techniques rely on an appropriately formulated criterion function to elicit the optimal subset of features from a larger feature set. In the case of a biometric system, this criterion function could be the Equal Error Rate (EER); the d-prime measure; the area of overlap between genuine and impostor training scores; or the average GAR at predetermined FAR values in the ROC/DET curves corresponding to the training set (see [1]).

Dimensionality reduction may also be accomplished using feature *transformation* methods where the vector $\boldsymbol{Z}'$ is subjected to a linear or a nonlinear mapping that projects it to a lower dimensional subspace. Examples of such transformations include the use of principal component analysis (PCA), independent component analysis (ICA), multidimensional scaling (MDS), Kohonen Maps, and neural networks [8]. The application of a feature selection or feature transformation procedure results in a new feature vector $\boldsymbol{Z} = \{z_1, z_2, \ldots z_k\}$ which can now be used to represent the identity of an individual.

Examples of Feature Level Fusion

Ross and Govindarajan [1] discuss feature level fusion as applied to three different scenarios: (1) multialgorithm, where two different face recognition algorithms based on Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are combined; (2) multisensor, where the three different color channels of a face image are independently subjected to LDA and then combined; and (3) multimodal, where the face and hand geometry feature vectors are combined. The general procedure adopted in [1] is summarized as follows.

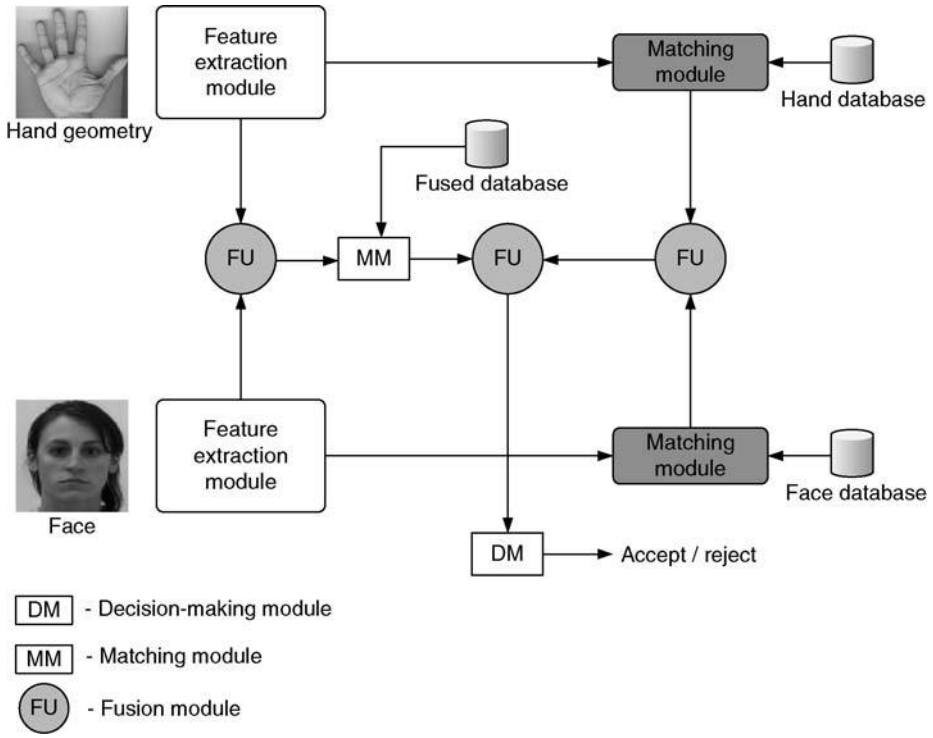
1. Let $\{\boldsymbol{X}_i, \boldsymbol{Y}_i\}$ and $\{\boldsymbol{X}_j, \boldsymbol{Y}_j\}$ be the feature vectors obtained at two different time instances i and j. Here, X and Y represent the feature vectors derived from two different information sources. The corresponding fused feature vectors may be denoted as $\boldsymbol{Z}_i$ and $\boldsymbol{Z}_j$, respectively.
2. Let s_X and s_Y be the normalized match scores generated by comparing $\boldsymbol{X}_i$ with $\boldsymbol{X}_j$ and $\boldsymbol{Y}_i$ with $\boldsymbol{Y}_j$, respectively, and let $s_{match} = (s_X + s_Y)/2$ be the fused match score obtained using the simple sum rule.
3. A pair of fused feature vectors, $\boldsymbol{Z}_i$ and $\boldsymbol{Z}_j$, are then compared using two different distance measures: the Euclidean distance (s_{euc}) and the Thresholded Absolute Distance or TAD (s_{tad}). Thus,

$$s_{euc} \sum_{r=1}^{k} (z_{i,r} - z_{j,r})^2 \tag{3}$$

$$s_{tad} \sum_{r=1}^{k} I(|z_{i,r} - z_{j,r}|, t). \tag{4}$$

   Here, $I(u,t) = 1$, if $u > t$ (and 0, otherwise), t is a prespecified threshold, and k is the dimensionality of the fused feature vector. The thresholded absolute distance measure determines the *number* of normalized feature values that differ by a magnitude greater than t. The s_{euc} and s_{tad} values are consolidated into one feature level score, s_{feat}, via the simple sum rule (Fig. 2). This retains information at the match score level (s_{match}) as well as the feature level (s_{feat}).
4. Finally, the simple sum rule is used to combine s_{match} and s_{feat} in order to obtain the final score s_{tot} (Fig. 3).

The authors compare the matching performances obtained using s_{match} and s_{tot} in all three scenarios. Results indicate that feature level fusion is advantageous in some cases. The feature selection scheme ensures that redundant or correlated feature values are detected and removed before invoking the matcher. This is probably one of the key benefits of performing fusion at the feature level [9].



Fusion, Feature-Level. Figure 3 The flow of information when data from the feature level and match score level are combined in a multibiometric system [1].

Chibelushi et al. [10] discuss a scheme to combine the features associated with the voice (audio) and lip shape (video) of an individual in an identification system. Fourteen mel-frequency cepstral coefficients (MFCC) and 12 geometric features are extracted from the audio and video streams to represent the voice and shape of the lips, respectively. The PCA and LDA transformations are used to reduce the dimensionality of the concatenated feature set. The authors demonstrate that the use of feature level fusion in their system is equivalent to increasing the signal-to-noise ratio (SNR) of the audio signal thereby justifying the use of lip shape in the fusion module. Other examples of feature level fusion can be found in [11] (face and iris) and [12] (hand geometry and palmprint).

Summary

Feature-level fusion represents an early fusion strategy in which multiple feature sets are consolidated in order to generate a more robust template. These feature sets can emerge (1) from a single biometric algorithm operating on different biometric samples (e.g., two images of the right hand of a single subject), or (2) from multiple biometric algorithms. If the feature sets to be combined originate from the same biometric algorithm (thus, a single modality is assumed), then feature level fusion can be used for template update or template improvement. If the feature sets originate from multiple biometric algorithms, then a concatenation procedure can be used to integrate them. The concatenation procedure has a feature normalization and a feature selection (or transformation) stage resulting in a compact set of salient features that can be used by the matcher. The primary advantage of such an approach is the elimination of redundant features thereby improving matching accuracy. In some cases, it may be advantageous to design a hybrid system that combines the outputs of score-level fusion and feature-level fusion. The disadvantages of feature-level fusion include the need to design a new matcher and to acquire a large number of training samples.

Related Entries

► Multibiometrics

References

1. Ross, A., Govindarajan, R.: Feature Level fusion using hand and face biometrics. In: Proceedings of SPIE Conference on Biometric Technology for Human Identification II. vol. 5779, pp. 196–204. Orlando, USA (2005)
2. Jiang, X., Ser, W.: Online fingerprint template improvement. IEEE Trans. Pattern Analy. Mach. Intell. **24**, 1121–1126 (2002)
3. Moon, Y.S., Yeung, H.W., Chan, K.C., Chan, S.O.: Template synthesis and image mosaicking for fingerprint registration: an experimental study. In: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). vol. 5, pp. 409–412. Montreal, Canada (2004)
4. Yau, W.Y., Toh, K.A., Jiang, X., Chen, T.P., Lu, J.: On fingerprint template synthesis. In: CD-ROM Proceedings of Sixth International Conference on Control, Automation, Robotics and Vision (ICARCV), Singapore (2000)
5. Jain, A.K., Chandrasekaran, B.: Dimensionality and sample size considerations in pattern recognition practice. In: Krishnaiah, P., Kanal L.N. (eds.) Handbook of Statistics, Vol. 2, Vol. 2, pp. 835–855. North-Holland, Amsterdam (1982)
6. Pudil, P., Novovicova, J., Kittler, J.: Floating search methods in feature selection. Pattern Recogn. Lett. **15**, 1119–1124 (1994)
7. Jain, A.K., Zongker, D.: Feature selection: evaluation, application, and small sample performance. IEEE Trans. Pattern Analy. Mach. Intell. **19**, 153–158 (1997)
8. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: a review. IEEE Trans. Pattern Analy. Mach. Intell. **22**, 4–37 (2000)
9. Kumar, A., Zhang, D.: Biometric recognition using feature selection and combination. In: Proceedings of Fifth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 813–822. Rye Brook, USA (2005)
10. Chibelushi, C.C., Mason, J.S.D., Deravi, F.: Feature-level data fusion for bimodal person recognition. In: Proceedings of the Sixth International Conference on Image Processing and Its Applications, vol. 1, pp. 399–403. Dublin, Ireland (1997)
11. Son, B., Lee, Y.: Biometric authentication system using reduced joint feature vector of iris and face. In: Proceedings of Fifth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 513–522. Rye Brook, USA (2005)
12. Kumar, A., Wong, D.C.M., Shen, H.C., Jain, A.K.: Personal Verification using palmprint and hand geometry biometric. In: Fourth International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA), pp. 668–678. Guildford, UK (2003)

Fusion, Image Level

► Fusion, Sensor-Level

Fusion, Measurement Level

► Fusion, Score-Level

Fusion, Physics-Based

Physics-based fusion makes use of the physical characteristics of the multispectral image acquisition process. In this fusion scheme, information on the spectral response of the sensor, the transmittance of the liquid crystal tunable filter (when used), the spectral reflectance of the object being imaged, and the spectral power distribution of the illuminant, are used, separately or in combination, as weights for the different sub-spectral images for their fusion.

► Multispectral and Hyperspectral Biometrics

Fusion, Quality-Based

Norman Poh
CVSSP, FEPS, University of Surrey Guildford, Surrey GU2 7XH, UK

Synonym

Quality-dependent fusion

Definition

Quality-based fusion refers to the use of *quality measures* in combining several biometric system outputs. Quality measures are an array of measurements quantifying the degree of excellence or conformance of biometric samples to some predefined criteria known to influence the system performance. Examples of quality measures for face biometrics are focus, contrast, and face detection reliability; and for iris biometrics are iris texture richness, the area of iris used for matching, and iris detection reliability. In quality-based fusion, the

match scores of biometric samples of higher quality are considered more important, i.e., given higher weights, in order to compute the final combined score.

Introduction

Quality-based fusion in the context of multibiometric systems is more challenging than multi-algorithmic systems because quality measures of the different biometric modalities are not comparable. This implies that quality-based fusion techniques have to necessarily consider the joint space of scores and quality measures, hence taking into account not only the dependency among scores themselves but also the dependency between scores and quality measures.

Prior studies in this direction include but are not limited to [1–4]. Nandakumar et al. proposed a likelihood ratio-based approach to achieve quality dependent score fusion [1]. This is a generative approach to model the relationship between scores and quality measures of the same modality. The likelihood of scores and quality measures of different biometric modalities are combined using the product rule, hence, realizing a naive Bayes classifier. The result is that the less informative modalities will produce likelihood ratios close to one and will therefore not influence the final combined score.

Fierrez-Aguilar et al. proposed a quality-based fusion realized using a support vector machine (SVM) [2]. In their context, quality measures were manually annotated and were used in two ways. First, they were used to control the penalty function of the SVM learning criterion. Second, during inference, quality measures were also used to weigh the relative influence of the respective modalities and the joint decision making process. Intuitively, the approach enables the multimodal system to focus on the single modality of dominant quality or for comparable qualities on the joint decision making system. Unfortunately, as a result of the SVM training strategy the joint decision making system is optimized for good quality data only.

Bigun et al. proposed the Bayesian Conciliation method [3]. This method relies on two components known as a client and an impostor supervisor. The client supervisor estimates the expected true authenticity score of a claim based on its expertise in recognizing client data (likewise for the impostor supervisor). The final decision is made by taking into account the different expertise of the two supervisors and choosing the one which comes closest to its goal, which is defined as zero for impostor supervisor and one for client supervisor. Effectively, the supervisor adapts to each identity claim as a function of the quality of the input data.

Kryszczuk et al. proposed a *derived* quality measure [4] instead of raw quality measures as done in [1–4]. The derived quality measure, or the confidence is defined as the posterior probability of making the correct decision given some observed evidences, which include both the system output and raw quality measures. In the context of bimodal fusion, this means that if the decision of two systems are in conflict (different), one takes the decision of the system which is more likely to be correct.

Kittler et al. proposed a framework to incorporate the quality information in fusion from a pattern recognition perspective [5]. In this framework, various levels of system output dependency, i.e., whether they belong to the same modality or to different modalities, are considered.

Last but not least, Poh et al. proposed a generative approach to estimate the joint density of scores and quality measures by first clustering the quality measures into discrete hidden states [6]. This approach assumes that the scores and quality measures are independent given the discrete quality state/cluster. This approach is sensible because similar quality measures in a cluster will share similar statistical properties and thus they can be combined by the same fusion classifier, and vice versa for dissimilar quality measures.

Quality-Based Fusion from the Pattern Recognition Perspective

Let $x \in \mathbb{R}^R$ be a vector of output scores of R experts, $q \in \mathbb{R}^P$ be a vector of P quality measures and $k \in \{C, I\}$ be one of the two possible classes of users, i.e., genuine users or clients and impostors. From the Bayesian point of view, the *generative* and *discriminative* approaches which incorporate the quality information directly can be written as follows:

$$y_{com}^{llr} \equiv f^{llr}(x,\ q) = \log \frac{p(x, q|C)}{p(x, q|I)} \qquad (1)$$

$$y_{com}^{prob} \equiv f^{prob}(x,q) = P(C|x,q) \tag{2}$$

In practice (2) is approximated by:

$$P(C|x,q) \approx \text{sigmoid}\left(f^{disc}(x,q)\right) = \frac{1}{1+\exp\left(f^{disc}(x,q)\right)} \tag{3}$$

where the output $f^{disc}(x,q) \in [-\infty,\infty]$ does not have to be associated with probability. $f^{disc}(x,q)$ is known as a discriminative function and very often, based on the sign of its output, one classifies x as either belonging to a client or an impostor. One can implement $f^{llr}(x,q)$ using any density estimator, e.g., Gaussian Mixture Model and Parzen windows [7]; $f^{prob}(x,q)$ using logistic regression [8] or any neural network [6] with the sigmoid activation function; and $f^{disc}(x,q)$ using a support vector machine [9], linear or quadratic discriminant functions and their variant [8], and neural networks.

The conventional fusion approaches without using the quality information can also be divided into either generative or discriminative. They can be written in similar ways as in (1) and (2) except that q is not used as part of the observations.

Classifier Design and System Output Dependency

Considered here is the case where the system outputs, x, can be obtained from the same biometric modality or from different modalities. For this reason, $x_{m,i}$ is introduced to denote the i-th classifier of the m-th biometric modality. There are I_m systems for the m-th modality and M biometric modalities are available. As a result, the number of systems available for fusion is $\sum_m I_m$.

In general, higher dependence is expected among the system outputs sharing the same biometric modality and, in contrast, independence when the system fuses different biometric modalities. By assuming different types of system output dependency, the following three types of fusion architecture are identified, in increasing levels of complexity:

1. *Multi-stage single processing (MSSP)*. This architecture is a result of assuming independence among all the system outputs despite the fact that systems sharing the same biometric modality may be dependent. It can be written as:

$$y_{com}^{MSSP} = \prod_m \prod_i P(C|x_{m,i},q) = \prod_m \prod_i f^{prob}(x_{m,i},q) \tag{4}$$

Note that since $f^{prob}(x_{m,i},q)$ operates on a single system at a time, it can be considered as a quality-dependent score normalization procedure. It is therefore not a deterministic one-to-one mapping function as studied in [10] but rather a function of $x_{m,i}$ and q jointly. Note that discriminative functions $f^{disc}(x,q)$, e.g., a Support Vector Machine (SVM), do not output scores which satisfy the axiomatic properties of probabilities and cannot therefore be used in conjunction with a product fusion rule. Instead, the sum rule may be more appropriate, i.e.,

$$y_{com}^{MSSP} = \sum_m \sum_i f^{disc}(x_{m,i},q) \tag{5}$$

By doing so, one implicitly assumes that the class-conditional distributions of the outputs $f^{disc}(x_{m,i},q)$ across all m and i are comparable. This is, in general, not the case, thus implying the need for normalizing the outputs. Fortunately, this can be avoided by normalizing the input to the function $f^{disc}: \mathbb{R}^{R+P} \to \mathbb{R}$ instead of its output. Suppose that each of the $\mathbb{R}^{R+P}$ input elements is normalized to having zero mean and unit variance (across all the training examples), and the same complexity of $f^{disc}(x_{m,i},q)$ is used for all m and i, then the outputs $f^{disc}(x_{m,i},q)$ will be comparable. For the generative approach, using the sum rule, i.e.,

$$y_{com}^{MSSP} = \sum_m \sum_i f^{llr}(x_{m,i},q), \tag{6}$$

is a direct implication of assuming independence among the output of systems $x_{m,i}$ for all m and i.

2. *Multi-stage joint processing (MSJP)*. This architecture takes into consideration the dependency among system outputs derived from the same biometric modality yet ignores the dependency of the system outputs coming from different biometric modalities. It can be written as:

$$y_{com}^{MSJP} = \prod_m P(C|x_m,q) = \prod_m f^{prob}(x_m,q), \tag{7}$$

where x_m denotes a vector the components of which are the system outputs sharing the m-th biometric modality, i.e., $x_m \equiv [x_{m,1},\ldots,x_{m,I_m}]$.

The practical implication of this architecture is that one designs a fusion classifier per biometric modality and then combines all M resulting fusion classifiers using a fixed rule, e.g., the product rule for $f^{prob}(x_m, q)$ and the sum rule for $f^{disc}(x_m, q)$ and $f^{llr}(x_m, q)$.

3. *Single-stage joint processing (SSJP)*. This architecture does not assume system output independence. It can be written as:

$$y_{com}^{SSJP} = P(C|x, q) = f^{prob}(x, q), \tag{8}$$

where x is a vector containing all the system outputs, i.e., $x = \{x_{m,i} | \forall_{i,m}\}$. The function $f^{prob}(x, q)$ is simply replaced by $f^{llr}(x, q)$ when using a ▶ generative classifier and by $f^{disc}(x, q)$ when using a ▶ discriminative classifier.

In the discussion that follows, the focus is on training the discriminative function $f^{disc}(x, q)$. However, the discussion generalizes to the functions $f^{llr}(x, q)$ and $f^{prob}(x, q)$. For this reason, the generic term $f(x, q)$ is used and refer to one of the three particular fusion algorithms, i.e., $f^{llr}(x, q)$, $f^{prob}(x, q)$, or $f^{disc}(x, q)$, only when necessary.

The Complexity of Modeling Scores and Quality Measures: A Generative Approach

In the generative approach, modeling the joint space of x and q is difficult since q is not directly relevant to the classification task. For example, if one uses a mixture of Gaussian components to estimate the joint density, one would use many more components than one does if one models just x. This problem is particularly acute when the dimension of q is large. One way to reduce the complexity (the number of components and their associated parameters) is to first cluster the quality measures and then learn the density of x for each cluster. This strategy was reported in [6]. Instead of modeling $p(x, q)$ directly, Poh et al. proposed to factorize it into $p(x|q)p(q)$ where,

$$p(x|q) = \sum_{Q} p(x|Q)P(Q|q) \tag{9}$$

where Q is a cluster state and $P(Q|q)$ is the posterior probability of Q given the observation q. Since Q is not observed (hidden), it has to be integrated out, hence, explaining the sum over Q in (9). In [6], it turns out that one does not need to model $p(q)$ to implement a quality-based fusion classifier.

The solution of (9) is more elegant than the one that directly estimates $p(x, q)$. This is because the density $p(x|Q)$ has only R dimensions, i.e., the dimension in x, whereas $p(x, q)$ has $R + P$ dimensions. As a result, one can potentially face the curse of dimensionality when modeling $p(x, q)$, especially in the situation where x is small and q is large in dimension. In brief, this curse means that modeling the increased number of dimensions may be less effective since this is not necessarily supported by an exponential increase in the number of training samples. In fact, there is only a fixed number of training samples to design one fusion classifier. Note that when q is one dimensional, the classifier should be more appropriately called a quality-dependent score normalization procedure.

The realized quality-based fusion via (9), when written in the form of (1), is:

$$f^{llr}(x, q) = \log \frac{\sum_Q p(x|\mathrm{C}, Q)p(Q|q)}{\sum_Q p(x|\mathrm{I}, Q)p(Q|q)} \tag{10}$$

The Complexity of Modeling Scores and Quality Measures: A Discriminative Approach

Similar to the generative approach, jointly estimating x and q is also a challenging problem for the discriminative approach. Suppose that, one uses a linear function in $f(x, q)$ to distinguish the client class from the impostor one. In this case a weight will be associated with each element in x and q. The result after training is that magnitude of the weight associated with q will be comparatively small because q has no discriminative information. This suggests that using nonlinear function of $f(x, q)$ may be more useful.

One way to introduce non-linearity is by using some kind of expansion between x and q, i.e., $x \otimes q$, where $\otimes$ is called a *tensor product*. Note that x and q are not vectors of the same length. If x has R elements and q has P elements, then $x \otimes q$ will result in $P \times R$ elements and each element is a product between a pair of the elements in x and q. Therefore, when training $f(x, q)$, the fusion classifier must be fed with inputs $[x, q, x \otimes q]$ instead of $[x, q]$.

When one uses $[x, q, x \otimes q]$, the linear function can be written as:

$$f(x,q) = \sum_i \sum_j w_{i,j} x_i q_j + \sum_i w_i x_i + \sum_j v_j q_j$$

$$= \sum_i x_i \underbrace{\left(\sum_j q_j w_{i,j} + w_i \right)}_{} + \underbrace{\sum_j v_j q_j}_{}, \quad (11)$$

where the weight $w_{i,j}$ is associated with $x_i q_j$, the weight w_i is associated with x_i, and v_j is associated with q_j. In this notation, x_i is an element of vector x and q_j is an element of vector q. (11) clearly shows that the resulting classifier is *linear* except that the weight is modified *dynamically* by the quality measures via the first under-braced term. The second under-braced term shows that q *dynamically* adjusts the decision threshold.

Several possible "arrangements" are outlined in Table 1, presented in the order of increasing complexity, i.e., the number of parameters. $f([x, q])$ is written to explicitly refer to the second arrangement, $f([x, x \otimes q])$ to refer to the third arrangement, etc. The second column shows the four possible arrangements, i.e., the way the features are used as input to a fusion algorithm. The third column shows the resulting discriminative linear function $f^{disc}(x, q)$. While similar analyzes cannot be done for the linear discriminative function $f^{prob}(x, q)$ (due to the sigmoid function) and for the generative function $f^{llr}(x, q)$, our purpose in showing the elements in the expanded input vector along with their associated weight parameters is to illustrate the complexity of each arrangement. For instance, the first arrangement, i.e., $f([x])$, does not use any quality information. The second arrangement, i.e., $f([x, q])$ does not contain any interaction between x and q. However, it considers the case where the decision threshold may be modified by q. In the third arrangement, one creates a linear classifier whose weights can change dynamically as a function of q. The last arrangement, i.e., $f([x, q, x \otimes q])$ or (11), is the most general one since it contains all possible interactions between x and q of the first three arrangements. In [5], it was shown that the last three arrangements achieve superior results compared to the first one (without considering the quality information) across many intramodal and multimodal fusion tasks.

Fusion, Quality-Based. Table 1 The complexity of the function $f(x,q)$ when implemented using a linear classifier, in increasing level of complexity due to different input arrangements

No.	Arrangement	The resulting function $f^{disc}(x,q)$	No. of parameters
1	$[x]$	$\sum_i x_i w_i$	R
2	$[x,q]$	$\sum_i x_i w_i + \sum_j q_j v_j$	$R+P$
3	$[x, x \otimes q]$	$\sum_i x_i \, (\sum_j q_j w_{i,j} + w_i)$	$R \times (P+1)$
4	$[x, q, x \otimes q]$	$\sum_i x_i$ $(\sum_j q_j w_{i,j} + w_i) + \sum_j v_j q_j$	$R+P+R \times P$

The quality-enhanced discriminative fusion classifier with the input $[x, x \otimes q]$ (the third arrangement) is structurally very similar to the one proposed in [11] where a reduced polynomial discriminative function was used. In our case, one can use *any* discriminative classifier to implement it. This is an elegant solution because one does not need to design a dedicated fusion algorithm such as those proposed in [2, 3, 11] to achieve the same goal any longer.

Related Entries

- ▶ Biometric Sample Quality
- ▶ Face Sample Quality
- ▶ Feature-level Fusion
- ▶ Fingerprint Image Quality
- ▶ Iris Image Quality
- ▶ Multiple Classifier Systems
- ▶ Multibiometrics
- ▶ Score-level Fusion

References

1. Nandakumar, K., Chen, Y., Dass, S., Jain, A.: Quality-based score level fusion in multibiometric systems. In: Proceedings of the 18th International Conference on Pattern Recognition (ICPR), pp. 473–476. Hong Kong (2006)
2. Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., Bigun, J.: Kernel-based multimodal biometric verification using quality signals. In: Defense and Security Symposium, Workshop on Biometric Technology for Human Identification, Proceedings of SPIE, vol. 5404, pp. 544–554 (2004)
3. Bigun, J., Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Multimodal biometric authentication using quality signals in mobile communications. In: 12th International Conference on Image Analysis and Processing, pp. 2–13. Mantova (2003)
4. Kryszczuk, K., Richiardi, J., Prodanov, P., Drygajlo, A.: Error handling in multimodal biometric systems using reliability

measures. In: Proceedings of the 12th European Conference on Signal Processing. Antalya, Turkey (2005)
5. Kittler, J., Poh, N., Fatukasi, O., Messer, K., Kryszczuk, K., Richiardi, J., Drygajlo, A.: Quality dependent fusion of intramodal and multimodal biometric experts. In: Proceedings of SPIE Defense and Security Symposium, Workshop on Biometric Technology for Human Identification, vol. 6539 (2007)
6. Poh, N., Heusch, G., Kittler, J.: On Combination of Face Authentication Experts by a Mixture of Quality Dependent Fusion Classifiers. In: LNCS 4472, Multiple Classifiers System (MCS), pp. 344–356. Prague (2007)
7. Bishop, C.: Neural Networks for Pattern Recognition. Oxford University Press (1999)
8. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer (2001)
9. Vapnik, V.N.: Statistical Learning Theory. Springer (1998)
10. Jain, A., Nandakumar, K., Ross, A.: Score normalisation in multimodal biometric systems. Pattern Recognit. **38**(12), 2270–2285 (2005)
11. Toh, K.A., Yau, W.Y., Lim, E., Chen, L., Ng., C.H.: Fusion of Auxiliary Information for Multimodal Biometric Authentication. In: LNCS 3072, International Conference on Biometric Authentication (ICBA), pp. 678–685. Hong Kong (2004)

Fusion, Rank-Level

Ajay Kumar
Department of Computing, The Hong Kong Polytechnic University

Synonym

Biometric Fusion, Rank-Level

Definition

Rank level fusion is the method of consolidating more than two identification results to enhance the reliability in personal identification. In multimodal biometric system, rank level fusion can be used to combine the biometrics matching scores from the different biometric modalities (for example face, fingerprint, palmprint, and iris). It can also be used for performance improvement in unimodal biometric system by combining multiple classifier output that use different classifiers (*K* nearest neighbor, neural network, support vector machine, decision tree, etc.), different training set, different architectures (different number of layers or transfer function in neural network), or different parameter values (different kernels in support vector machine or different *K* in *K* nearest neighbor).

Introduction

The majority of biometric system deployed using feature extraction from a single biometric modality and a particular classification procedure to determine the identity on an individual. The perfect solutions for user identification are often difficult to achieve, mainly due to the large number of user classes and the imperfection in the feature extraction process. Therefore, the improvement in the user identification results using the simultaneous extraction of features and classifiers of different types has been investigated. The combination of potentially conflicting decisions in multimodal or unimodal biometric system employing different classifiers can be achieved in several ways: at feature, score, and decision level. In general, the improvement in identification accuracy is achieved by selecting combination mechanism that can take advantage of strengths of individual classifiers while suppressing their weakness.

Any biometric recognition system is capable of generating matching scores for the input user with those of the enrolled possible identities. The set of all the possible user identities can be ranked by sorting the matching scores in the descending order. Thus a biometric system can identify an unknown user by generating ranks, i.e., integer numbers for each of the possible user identity. The rank level fusion refers to the mechanism of combining such output ranks from the various biometrics ▶ matchers (subsystems), to consolidate the combined output ranks to establish the identity of an individual with higher confidence. The matching score contains more information than ranks and therefore matching score level fusion schemes are believed to be more flexible. However, the rank level fusion schemes do not require ▶ transformation of ranks from various biometrics matchers into a common domain and are simpler to implement. Several decision level fusion schemes only use ▶ top choice (rank) from each of the biometric classifiers, which is likely to be sufficient for biometric systems with small number of users. However, with the increase in number of enrolled identities or users, the correct rate for top choices drops, the ▶ secondary choices often contain near misses that should not be overlooked and are made use of in the rank level fusion.

Methods for Combining Ranks

The voting techniques proposed by different researchers [1–3] for consolidating rank output from the different biometric matchers will now be introduced. Given the ranked list of user identities returned by M different biometric matchers, let $r_i(k)$ be the rank assigned to the user k by the ith matcher. The user identity for kth user is assigned by computing the fused rank score m_k from all the M matchers.

1. *The Highest Rank Method.* In this method, the user identity is ascertained from the highest ranks returned by the individual matchers. Each of the possible user identity receives M ranks, each from the M matchers. The fused rank score m_k for every possible user identity k is computed from the minimum (highest) of these M ranks. The user identities are then sorted in the order of fused rank scores to obtain the combined or new ranking from all M matchers. Any ties in the fused rank scores (m_k) are randomly broken to obtain linearly ordered combined ranking. These ties are due to a number of user identities sharing the same combined ranks and depend on the number of employed matchers. The chances of the occurrences of such ties will be smaller, if the number of enrolled user identities are large and the number of matchers employed in the fusion are small. The advantage of this method lies in the utilization of strength of each of the biometric matchers. However, large number of matchers can result in more ties in the combined ranking, which is the major problem in this method. Therefore this method is considered useful in biometric systems combining small number of matchers with large number of enrolled users.
2. *Borda* (Named for the French scientist *Jean-Charles de Borda* (1733–1799) who formulated this preferential voting system.) *Count Method.* The Borda count is the generalization of majority vote and the most commonly used method for ► unsupervised rank level fusion. It is the voting method in which each matcher gives priority to all possible user identities. Each matcher ranks the fixed set of possible user identities in the order of its preference. For every matcher, the top ranked user identity is given N votes, the second ranked candidate identity is given N-1 votes and so on. Then for every possible user identity, the votes from all the matchers are added. The user identity that receives the highest number of votes is assigned as the winner or the true user identity.

$$m_k = \sum_{i=1}^{M} r_i(k) \, \forall k, \quad k = \{1, 2, \ldots N\}. \tag{1}$$

The Borda count score m_k represents strength of agreement among different biometric matchers. The Borda count method assumes statistical independence, i.e., ranks assigned to a given user by different matchers are independent. This assumption is often made in practice but it may not be true. The Borda count method is particularly considered suitable for combining the biometrics matchers with large number of user identities that often generate the correct user identities *near* the top of list (ranks) but not *at* the top. This method is efficient, simple, and does not require any training. However, it assumes that all matchers are equally correct. This may not be the case when some matchers are more likely to be correct than others. Therefore, weighted Borda count method has been suggested to utilize the strength of individual matchers.

3. *Weighted Borda Count Method.* The performance of different biometric matchers is not uniform, for example, a biometric matcher using iris images is expected to perform better than those matchers using hand geometry or face images. Therefore, modification of Borda count method by assigning corresponding weights to the ranks produced by individual matchers has been suggested. The fused rank scores in weighted Borda count method are computed as follows:

$$m_k = \sum_{i=1}^{M} w_i r_i(k), \tag{2}$$

where the w_i represents the weights assigned to the ith matcher. The weight w_i are assigned to reflect the significant of each matcher and can be computed from the overall assessment of the performance. The weights are computed during the training phase using logistic regression (as detailed in [3]) or using more sophisticated machine learning techniques.

4. *Bayes Fuse.* The Bayes fuse is the ► supervised rank level fusion method based on Bayesian inference. Each of the possible user identity is ranked

according to the fused rank scores computed as follows:

$$m_k = \sum_{i=1}^{M} \log \frac{\Pr[m_k(i)|genuine]}{\Pr[m_k(i)|imposter]}, \qquad (3)$$

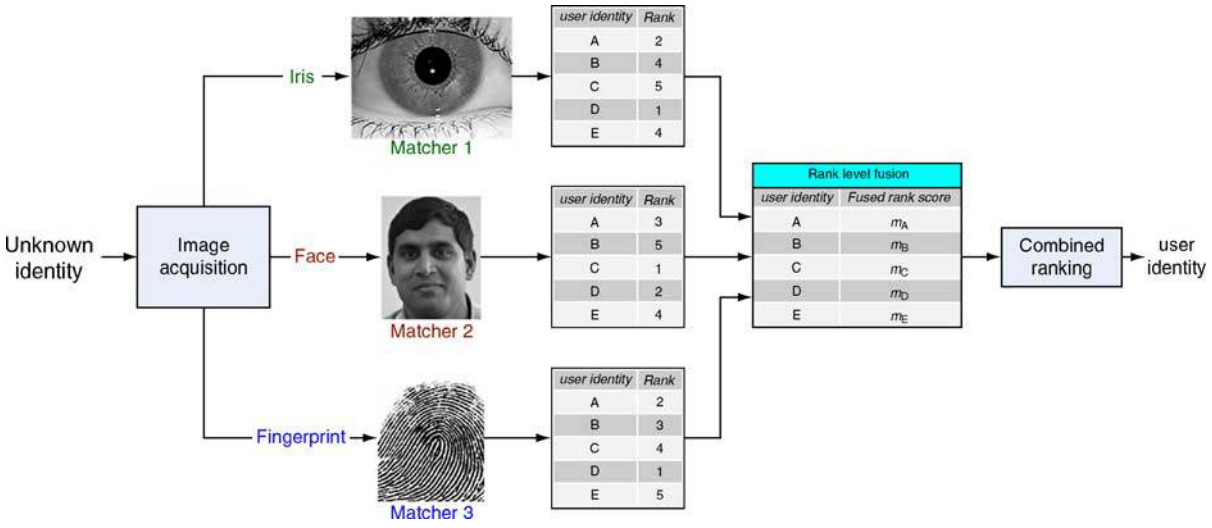
where $P_r[m_k(i)|\ imposter]$ is the probability that an imposter user would be ranked to $m_k(i)$ by the ith matcher and $P_r[m_k(i)|\ genuine]$ is the probability that a genuine user would be ranked to $m_k(i)$ by the ith matcher. These two likelihood probabilities are computed from the training data during training phase. The above equation is easily derived [2] from the estimation of two posterior probabilities, each for the genuine and imposter class, using Bayes rule. The combined ranks generated using Eq. (3) makes a common naive Bayes assumption, i.e., individual ranks assigned to the user identities by M matchers are independent. The training phase in Bayes fuse method required the collection of simple statistics about the distribution of ranks among various user identities. The rank level fusion using Bayes fuse was originally introduced for information retrieval but is equally useful in biometrics fusion.

Example

The four different rank level fusion methods discussed above can be better clarified with a simple example in multimodal biometric fusion. This example illustrates the combination of three different biometric matchers (Fig. 1), using iris, fingerprint, and face image, to generate matching scores. These matching scores are internally sorted to produce different ranking among the possible user identities. There are only five different users (user A, user B, user C, user D, and user E) and 1, 2, . . .5 represents the ranks for the possible input user identity with 1 being the highest rank/possibility. Let the weights of different matchers computed from the training data using linear regression be 0.5, 0.15, 0.35 for the matcher 1, matcher 2, and matcher 3 respectively. Let the probability that a genuine user be ranked at ranks (1, 2, 3, 4, 5) be (0.8, 0.1, 0.06, 0.02, 0.02), (0.5, 0.42, 0.06, 0.01, 0.01), and (0.6, 0.2, 0.08, 0.07, 0.05) for matcher 1, matcher 2, and matcher 3 respectively. Similarly the prior probabilities for an imposter user be ranked at ranks (1, 2, 3, 4, 5) have been obtained from the training data and are listed as (0.2, 0.9, 0.94, 0.98, 0.98), (0.5, 0.58, 0.94, 0.99, 0.99), and (0.4, 0.8, 0.92, 0.93, 0.95) respectively for matcher 1, matcher 2, and matcher 3.

Let us now compute the fused rank scores (m_A, m_B, m_C, m_D, m_E) and the new rankings for each of the four methods discussed in previous section.

1. *Highest Rank.* The fused rank scores using highest rank level method are shown in Table 2. The fused rank score for user A (m_A) will be 2 (highest rank or minimum of 2, 3, 2). The ties for m_C and m_D are randomly broken and the combined ranking is also shown in Table 1. The highest rank method achieves highest ranking for C and therefore the unknown input identity is user C.



Fusion, Rank-Level. **Figure 1** An example of multimodal biometric system employing rank level fusion.

Fusion, Rank-Level. **Table 1** Example for consolidating ranks using unsupervised rank level fusion methods

User identity	Highest rank method			Borda count method		
	Fused rank score		Combined ranking	Fused rank score		Combined ranking
A	m_A	2	3	m_A	7	2
B	m_B	3	4	m_B	12	4
C	m_C	1	1	m_C	10	3
D	m_D	1	2	m_D	4	1
E	m_E	4	5	m_E	13	5

Fusion, Rank-Level. **Table 2** Example for consolidating ranks using supervised rank level fusion methods

User identity	Weighted borda count method			Bayes fuse method		
	Fused rank score		Combined ranking	Fused rank score		Combined ranking
A	m_A	2.15	2	m_A	−6.34	2
B	m_B	3.8	3	m_B	−10.93	4
C	m_C	4.05	4	m_C	−6.48	3
D	m_D	1.15	1	m_D	1.47	1
E	m_E	4.35	5	m_E	−11.43	5

2. *Borda Count.* The fused rank scores using Borda count are computed as follows: $m_A = (2 + 3 + 2) = 7$, $m_B = (4 + 5 + 3) = 12$, $m_C = (5 + 1 + 4) = 10$, $m_D = (1 + 2 + 1) = 4$, $m_E = (4 + 4 + 5) = 13$. Thus, m_D is lowest and user D achieves highest combined ranking (Table 1).
3. *Weighted Borda Count.* The fused rank scores $m_A = (2 \times 0.5 + 3 \times 0.15 + 2 \times 0.35) = 2.15$. Similarly rank fused scores for rest of the users can be computed and are shown in Table 2.
4. *Bayes Fuse.* The prior probabilities that each of the ranks are true (untrue), i.e., belongs to the genuine (imposter) class, can be obtained from the training data and are provided in the problem. The fused rank score for user A can be computed using (3) as follows: $m_A = \log(0.1/0.9) + \log(0.06/0.94) + \log(0.2/0.8) = -6.34$. The rest of the fused rank scores and the combined rankings are displayed in Table 2.

Summary

In the biometrics literatures, one can find several examples [1, 3, 4, 6] of above rank level fusion methods to consolidate the outputs from different matchers. Bhatnagar et al. [4] employs a variation of Borda count method that uses partitioning of templates to consolidate the combined ranks. Highest rank method employed by Rautiainen and Seppanen [6], is referred as lowest rank method since it chooses the minimum rank from the list of dissimilarity score instead of conventional maximum rank methods that employ highest ranks from the list of similarity scores. Several other variations of Borda count method have also been developed in the literature [7]; *Nenson's method* that uses successive elimination from Borda count that are below average Borda count or *Quota Borda method* that includes the quota element in counting ranks. However, they have not yet been investigated for their utility in the biometrics literature.

A survey of biometrics on various fusion techniques [5] suggests that the rank level fusion method is less preferred method of fusion while score level fusion continues to be the most popular method. The rank level fusion can be more useful in combining decisions from a large number of biometric matchers and such large systems has not yet been evaluated in the biometrics literature.

References

1. Lee, Y., Lee, K., Jee, H., Gil, Y., Choi, W., Ahn, D., Pan, S.: "Fusion for multimodal biometric identification," Proc. ABVPA 2005, LNCS 3546, 1071–1079 (2005)

2. Aslam, J.A., Montague, M.: "Models for metasearch," In: Proceedings of the 24th ACM SIGIR Conference on Research and Development in Information Retrieval, Sep. 2001, pp. 379–381 (2001)
3. Ho, T.K., Hull, J.J., Srihari, N.: "Decision combination in multiple classifier systems." IEEE Trans. Pattern Anal. Mach. Intell. **16**, 66–75 (1994)
4. Bhatnagar, J., Kumar, A., Saggar, N.: "A novel approach to improve biometric recognition using rank level fusion," Proc. CVPR 2007, Minneapolis, MN, pp. 1–6, (2007)
5. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics, Springer Verlag (2006)
6. Rautiainen, M. and Seppanen, T.: "Comparison of visual features and fusion techniques in automatic detection of concepts from news video," In: Proceedings of the IEEE International Conference on Multimedia and Expo., ICME 2005, Amsterdam, pp. 932–935 (2005)
7. Dummett, M.A.E.: Principles of Electoral Reform, Oxford University, New York (1997)

Fusion, Score-Level

Arun Ross[1], Karthik Nandakumar[2]
[1]West Virginia University, Morgantown, WV, USA
[2]Institute for Infocomm Research A* STAR, Fusionopolis, Singapore

Synonyms

Fusion at the confidence level; Fusion at the measurement level; Match score fusion

Definition

In score-level fusion the match scores output by multiple biometric matchers are consolidated in order to render a decision about the identity of an individual. Typically, this consolidation procedure results in the generation of a single scalar score which is subsequently used by the biometric system. Fusion at this level is the most commonly discussed approach in the biometric literature primarily due to the ease of accessing and processing match scores (compared with the raw biometric data or the feature set extracted from the data). Fusion methods at this level can be broadly classified into three categories: density-based schemes, transformation-based schemes and classifier-based schemes.

Introduction

A match score is the result of comparing two feature sets extracted using the same feature extractor. A *similarity* score denotes how "similar" the two feature sets are, while a *distance* score denotes how "different" they are. Consequently, a high similarity score between a pair of feature sets indicates a good match whereas a high distance score indicates a poor match.

In score-level fusion the match scores output by multiple biometric matchers are combined to generate a new match score (a scalar) that can be subsequently used by the verification or identification modules for rendering an identity decision (alternatively, the fusion process may directly result in a decision). Fusion at this level is the most commonly discussed approach in the biometric literature primarily due to the ease of accessing and processing match scores (compared with the raw biometric data or the feature set extracted from the data). Fusion methods at this level can be broadly classified into three categories [1]: density-based schemes, transformation-based schemes and classifier-based schemes.

Density-Based Fusion schemes

Let $\mathbf{s} = [s_1, s_2, \ldots, s_R]$ denote the scores emitted by multiple matchers, with s_j representing the match score of the *jth* matcher, $j = 1, \ldots, R$. Further, let the labels ω_0 and ω_1 denote the genuine and impostor classes, respectively. Then, by ▶ Bayes decision theory [2], the probability of error can be minimized by adopting the following decision rule. (This is known as the Bayes decision rule or the minimum-error-rate classification rule under the 0-1 loss function [2]).

$$\text{Assign } \mathbf{s} \rightarrow \omega_i \text{ if}$$

$$P(\omega_i|\mathbf{s}) > P(\omega_j|\mathbf{s}), i \neq j, \quad \textit{and} \quad i,\ j = 0, 1. \qquad (1)$$

Here, the *a posteriori* probability $P(\omega_i|\mathbf{s})$, $i = 0,1$, can be derived from the class-conditional density function $p(\mathbf{s}|\omega_i)$ using the Bayes formula, i.e.,

$$P(\omega_i|\mathbf{s}) = \frac{p(\mathbf{s}|\omega_i)P(\omega_i)}{p(\mathbf{s})}, \qquad (2)$$

where $P(\omega_i)$ is the *a priori* probability of observing class ω_i and $p(\mathbf{s})$ denotes the probability of encountering $\mathbf{s}$. Thus, Eq. (1) can be re-written as

Assign $\mathbf{s} \rightarrow \omega_i$ if

$$\frac{p(\mathbf{s}|\omega_i)}{p(\mathbf{s}|\omega_j)} > \tau, i \neq j, \quad and \quad i, j = 0, 1 \tag{3}$$

where $\frac{p(\mathbf{s}|\omega_i)}{p(\mathbf{s}|\omega_j)}$ is known as the *likelihood ratio* and $\tau = \frac{P(\omega_j)}{P(\omega_i)}$ is a predetermined threshold. The density $p(\mathbf{s}|\omega_i)$ is typically estimated from a training set of match score vectors, using parametric or nonparametric techniques. However, a large number of training samples are necessary to reliably estimate the joint-density function $p(\mathbf{s}|\omega_i)$ especially if the dimensionality of the feature vector $\mathbf{s}$ is large. In the absence of sufficient number of training samples (which is typically the case when the multibiometric system is first deployed or if its parameters are subsequently adjusted), it is commonly assumed that the scalar scores $s_1, s_2, \ldots s_R$ are generated by R independent random processes. This assumption permits the density function to be expressed as

$$p(\mathbf{s}|\omega_i) = \prod_{j=1}^{R} p(s_j|\omega_i), \tag{4}$$

where the joint-density function is now replaced by the product of its marginals. The marginal densities, $p(s_j|\omega_i)$, $j = 1, 2, \ldots R$, $i = 0, 1$, are estimated from a training set of genuine and impostor scores corresponding to each of the R biometric matchers. Equation (4) results in the *product rule* which combines the scores generated by the R matchers as,

$$s_{prod} = \prod_{j=1}^{R} \frac{p(s_j|\omega_0)}{p(s_j|\omega_1)}. \tag{5}$$

Kittler et al. [3] modify the product rule by further assuming that the *a posteriori* probability $P(\omega_i|\mathbf{s})$ of class ω_i does not deviate much from its *a priori* probability $P(\omega_i)$ resulting in the *sum rule*:

$$s_{sum} = \frac{\sum_{j=1}^{R} p(s_j|\omega_0)}{\sum_{j=1}^{R} p(s_j|\omega_1)}. \tag{6}$$

Similar expressions can be derived for combining the match scores using the max, min, and median rules [1, 3]. All the aforementioned rules implicitly assume that the match scores are *continuous* random variables. Dass et al. [4] relax this assumption and represent the univariate density functions (i.e., the marginals in Eq. (4)) as a mixture of discrete as well as continuous components. The resulting density functions are referred to as generalized densities. The authors demonstrate that the use of generalized density estimates (as opposed to continuous density estimates) significantly enhances the matching performance of the fusion algorithm. Furthermore, they use ▶ copula functions to model the correlation structure between the match scores $s_1, s_2, \ldots, s_R$ and, subsequently, define a novel fusion rule known as the copula fusion rule.

Transformation-Based Fusion schemes

Density-based schemes, as stated earlier, require a large number of training samples (i.e., genuine and impostor match scores) in order to accurately estimate the density functions. This may not be possible in most multibiometric systems due to the time, effort, and cost involved in acquiring labeled multibiometric data in an operational environment. In such situations, it may be necessary to *directly* combine the match scores generated by multiple matchers using simple fusion operators (such as the simple sum of scores or order statistics) without first interpreting them in a probabilistic framework. However, such an approach is meaningful only when the scores output by the matchers are comparable. To facilitate this, a score normalization process is essential to transform the multiple match scores into a common domain (it must be noted, however, that some score normalization schemes do require a large number of training samples as seen in the following section). The process of score normalization entails changing the location and the scale parameters of the underlying match score distributions in order to ensure compatibility between multiple score variables. A few of the commonly discussed score normalization methods are described in this article.

The simplest normalization technique is the *min–max* normalization. Min–max normalization is best suited for the case where the bounds (maximum and minimum values) of the scores produced by a matcher are known. In this case, the minimum and maximum scores can be easily transformed into 0 and 1, respectively. However, even if the match scores are not bounded, the minimum and maximum values for the given set of training match scores can be estimated prior to applying min–max normalization. Let s_j^i denote the *ith* match score output by the *jth* matcher,

$i = 1, 2, \ldots, N$; $j = 1, 2, \ldots, R$ (R is the number of matchers and N is the number of match scores available in the training set). The min–max normalized score, ns_j^t, for the test score s_j^t is given by

$$ns_j^t = \frac{s_j^t - \min_{i=1}^{N} s_j^i}{\max_{i=1}^{N} s_j^i - \min_{i=1}^{N} s_j^i}. \quad (7)$$

When the minimum and maximum values are estimated from the given set of match scores, this method is not robust (i.e., the method is sensitive to outliers in the data used for estimation). Min–max normalization retains the original distribution of scores except for a scaling factor and transforms all the scores into a common range [0, 1]. Distance scores can be transformed into similarity scores by subtracting the normalized score from 1.

Decimal scaling can be applied when the scores of different matchers are on a logarithmic scale. For example, if one matcher has scores in the range [0, 10] and the other has scores in the range [0, 1000], the following normalization could be applied to transform the scores of both the matchers to the common [0, 1] range.

$$ns_j^t = \frac{s_j^t}{10^{n_j}}, \quad (8)$$

where $n_j = \log_{10} \max_{i=1}^{N} s_j^i$. In the example with two matchers where the score ranges are [0, 10] and [0, 1000], the values of n would be 1 and 3, respectively. The problems with this approach are the lack of robustness and the implicit assumption that the scores of different matchers vary by a logarithmic factor.

The most commonly used score normalization technique is the *z-score* normalization that uses the arithmetic mean and standard deviation of the training data. This scheme can be expected to perform well if the average and the variance of the score distributions of the matchers are available. If the values of these two parameters are not known, then they can be estimated based on the given training set. The z-score normalized score is given by

$$ns_j^t = \frac{s_j^t - \mu_j}{\sigma_j}, \quad (9)$$

where μ_j is the arithmetic mean and σ_j is the standard deviation for the *jth* matcher. However, both mean and standard deviation are sensitive to outliers and hence, this method is not robust. Z-score normalization does not guarantee a common numerical range for the normalized scores of the different matchers. If the distribution of the scores is not Gaussian, z-score normalization does not preserve the distribution of the given set of scores. This is due to the fact that mean and standard deviation are the optimal location and scale parameters only for a Gaussian distribution. While mean and standard deviation are reasonable estimates of location and scale, respectively, they are not optimal for an arbitrary match score distribution.

The *median* and *median absolute deviation* (MAD) statistics are less sensitive to outliers as well as points in the extreme tails of the distribution. Hence, a normalization scheme using median and MAD would be relatively robust and is given by

$$ns_j^t = \frac{s_j^t - med_j}{MAD_j}, \quad (10)$$

where $med_j = median_{i=1}^{N} s_j^i$ and $MAD_j = median_{i=1}^{N} |s_j^i - med_j|$. However, the median and the MAD estimators have a low efficiency compared to the mean and the standard deviation estimators, i.e., when the score distribution is not Gaussian, median and MAD are poor estimates of the location and scale parameters. Therefore, this normalization technique does not preserve the input score distribution and does not transform the scores into a common numerical range.

Cappelli et al. [5] use a *double sigmoid function* for score normalization in a multibiometric system that combines different fingerprint matchers. The normalized score is given by

$$ns_j^t = \begin{cases} \frac{1}{1+\exp\left(-2\left(\frac{s_j^t - \tau}{\alpha_1}\right)\right)} & \text{if } s_j^t < \tau, \\ \frac{1}{1+\exp\left(-2\left(\frac{s_j^t - \tau}{\alpha_2}\right)\right)} & \text{otherwise,} \end{cases} \quad (11)$$

where τ is the reference operating point and α_1 and α_2 denote the left and right edges of the region in which the function is linear. The double sigmoid function exhibits linear characteristics in the interval $(\tau - \alpha_1, \tau - \alpha_2)$. While the double sigmoid normalization scheme transforms the scores into the [0, 1] interval, it requires careful tuning of the parameters τ, α_1 and α_2 to obtain good efficiency. Generally, τ is chosen to be some value falling in the region of overlap between the genuine and impostor score distributions, and α_1 and α_2 are set so that they correspond to the

extent of overlap between the two distributions toward the left and right of τ, respectively. This normalization scheme provides a linear transformation of the scores in the region of overlap, while the scores outside this region are transformed nonlinearly. The double sigmoid normalization is very similar to the min–max normalization followed by the application of a two-quadrics (QQ) or a logistic (LG) function as suggested by [6]. When the values of α_1 and α_2 are large, the double sigmoid normalization closely resembles the QQ-min–max normalization. On the other hand, the double sigmoid normalization can be made to approach the LG-min–max normalization by assigning small values to α_1 and α_2.

The *tanh-estimators* introduced by Hampel [7] are robust and highly efficient. The tanh normalization is given by

$$ns_j^t = \frac{1}{2}\left\{\tanh\left(0.01\left(\frac{s_j^t - \mu_{GH}}{\sigma_{GH}}\right)\right) + 1\right\}, \qquad (12)$$

where μ_{GH} and σ_{GH} are the mean and standard deviation estimates, respectively, of the genuine score distribution as given by Hampel estimators. Hampel estimators are based on the following influence (ψ)-function:

$$\psi(u) = \begin{cases} u & 0 \le |u| < a, \\ a * sign(u) & a \le |u| < b, \\ a * sign(u) * \left(\frac{c-|u|}{c-b}\right) & b \le |u| < c, \\ 0 & |u| \ge c, \end{cases} \qquad (13)$$

where

$$sign\{u\} = \begin{cases} +1, & \text{if } u \ge 0, \\ -1, & \text{otherwise.} \end{cases} \qquad (14)$$

The Hampel influence function reduces the influence of the scores at the tails of the distribution (identified by *a*, *b*, and *c*) during the estimation of the location and scale parameters. Hence, this method is not sensitive to outliers. If many of the points that constitute the tail of the distributions are discarded, the estimate is robust but not efficient (optimal). On the other hand, if all the points that constitute the tail of the distributions are considered, the estimate is not robust but its efficiency increases. Therefore, the parameters a, b, and c must be carefully chosen depending on the amount of robustness required which in turn depends on the amount of noise in the available training data.

Fusion, Score-Level. Table 1 Summary of score normalization techniques.

Normalization technique	Robustness	Efficiency
Min–max	No	High
Decimal scaling	No	High
Z-score	No	High
Median and MAD	Yes	Moderate
Double sigmoid	Yes	High
Tanh-estimators	Yes	High

Mosteller and Tukey [8] introduce the biweight location and scale estimators that are robust and efficient. But, the *biweight estimators* are iterative in nature (initial estimates of the biweight location and scale parameters are chosen, and these estimates are updated based on the training scores), and are applicable only for Gaussian data. A summary of the characteristics of the different normalization techniques discussed in this article is shown in Table 1. The min–max, decimal scaling and z-score normalization schemes are efficient, but are not robust to outliers. On the other hand, the median normalization scheme is robust but inefficient. Only the double sigmoid and tanh-estimators have both the desired characteristics, namely, robustness and efficiency.

Once the match scores output by multiple matchers are transformed into a common domain they can be combined using simple fusion operators such as the sum of scores, product of scores or order statistics (e.g., maximum/minimum of scores or median score).

Classifier-Based Fusion schemes

In the verification mode of operation, the match scores generated by the multiple matchers may be input to a trained pattern classifier, such as a neural network, in order to determine the class label (genuine or impostor). In this approach, the goal is to directly estimate the class rather than to compute an intermediate scalar value. Classifier-based fusion schemes assume the availability of a large representative number of genuine and impostor scores during the training phase of the classifier when its parameters are computed. The component scores do not have to be transformed into a common domain prior to invoking the classifier.

In the biometric literature several classifiers have been used to consolidate the match scores of multiple matchers. Brunelli and Falavigna [9] use a HyperBF network to combine matchers based on voice and face features. Verlinde and Cholet [10] compare the relative performance of three different classifiers, namely, the k-Nearest Neighbor classifier using vector quantization, the decision tree classifier, and a classifier based on the logistic regression model while fusing the match scores originating from three biometric matchers. Experiments on the M2VTS database show that the total error rate (sum of the false accept and false reject rates) of the multimodal system is an order of magnitude less than that of the individual matchers. Chatzis et al. [11] use classical k-means clustering, fuzzy clustering and median radial basis function (MRBF) algorithms for fusion at the match score level. The proposed system combines the output of five different face and voice matchers. Each matcher provides a match score and a quality metric indicating the reliability of the match score. These values are concatenated to form a ten-dimensional vector that is input to the classifiers. Ben-Yacoub et al. [12] evaluate a number of classification schemes for fusion including support vector machine (SVM) with polynomial kernels, SVM with Gaussian kernels, C4.5 decision trees, multilayer perceptron, Fisher linear discriminant, and Bayesian classifier. Experimental evaluations on the XM2VTS database consisting of 295 subjects suggest the benefit of score level fusion. Bigun et al. [13] propose a novel algorithm based on the Bayesian classifier that takes into account the estimated accuracy of the individual classifiers (i.e., matchers) during the fusion process. Sanderson and Paliwal [14] use a support vector machine (SVM) to combine the scores of face and speech experts. In order to address noisy input, they design structurally noise-resistant classifiers based on a piece-wise linear classifier and a modified Bayesian classifier.

Summary

In a multibiometric system, fusion at the score level offers the best tradeoff between amount of information that is available and ease of fusion. Hence, score level fusion is typically adopted by most multibiometric systems. Although a wide variety of score level fusion techniques have been proposed in the literature, these can be grouped into three main categories, viz., density-based, transformation-based and classifier-based schemes. The performance of each scheme depends on the amount and quality of the available training data. If a large number of match scores is available for training the fusion module, then density-based approaches such as the likelihood ratio test can be used. Estimating the genuine and impostor distributions may not always be feasible due to the limited number of training samples that are available. In such cases, transformation-based schemes are a viable alternative. The nonhomogeneity of the match scores presented by the different matchers raises a number of challenges. Suitable score normalization schemes are essential in order to transform these match scores into a comparable domain. The sum of scores fusion method with simple score normalization (such as min–max) represents a commonly used transformation-based scheme. Classification-based fusion schemes consolidate the outputs of different matchers into a single vector of scores which is input to a trained classifier. The classifier then determines if this vector belongs to the "genuine" or "impostor" class.

Related Entries

- ▶ Fusion, Quality-Based
- ▶ Fusion, User-Specific
- ▶ Multibiometrics

References

1. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. 1st edn. Springer, New York, USA (2006)
2. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley, New York (2001)
3. Kittler, J., Hatef, M., Duin, R.P., Matas, J.G.: On combining classifiers. IEEE Trans. Pattern Analy. Mach. Intell. **20**, 226–239 (1998)
4. Dass, S.C., Nandakumar, K., Jain, A.K.: A principled approach to score level fusion in multimodal biometric systems. In: Proceedings of Fifth International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA), pp. 1049–1058. Rye Brook, USA (2005)
5. Cappelli, R., Maio, D., Maltoni, D.: Combining fingerprint classifiers. In: Proceedings of First International Workshop on Multiple Classifier Systems, pp. 351–361. Cagliari, Italy (2000)

6. Snelick, R., Uludag, U., Mink, A., Indovina, M., Jain, A.K.: Large Scale evaluation of multimodal biometric authentication using state-of-the-art systems. IEEE Trans. Pattern Analy. Mach. Intell. **27**, 450–455 (2005)
7. Hampel, F.R., Rousseeuw, P.J., Ronchetti, E.M., Stahel, W.A.: Robust Statistics: The Approach Based on Influence Functions. Wiley, New York (1986)
8. Mosteller, F., Tukey, J.W.: Data Analysis and Regression: A Second Course in Statistics. Addison-Wesley, Reading, MA, USA (1977)
9. Brunelli, R., Falavigna, D.: Person Identification using multiple cues. IEEE Trans. Pattern Analy. Mach. Intell. **17**, 955–966 (1995)
10. Verlinde, P., Cholet, G.: Comparing decision fusion paradigms using k-NN based classifiers, decision trees and logistic regression in a multi-modal identity verification application. In: Proceedings of Second International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 188–193. Washington D.C., USA (1999)
11. Chatzis, V., Bors, A.G., Pitas, I.: Multimodal decision-level fusion for person authentication. IEEE Trans. Syst. Man Cybernet. Part A: Syst. Humans **29**, 674–681 (1999)
12. Ben-Yacoub, S., Abdeljaoued, Y., Mayoraz, E.: Fusion of face and speech data for person identity verification. IEEE Trans. Neural Networks **10**, 1065–1075 (1999)
13. Bigun, E.S., Bigun, J., Duc, B., Fischer, S.: Expert Conciliation for multimodal person authentication systems using bayesian statistics. In: First International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA), pp. 291–300. Crans-Montana, Switzerland (1997)
14. Sanderson, C., Paliwal, K.K.: Information Fusion and Person Verification Using Speech and Face Information. Research Paper IDIAP-RR 02-33, IDIAP (2002)

Fusion, Sensor-Level

Afzel Noore, Richa Singh, Mayank Vasta
West Virginia University, Morgantown, WV, USA

Synonyms

Fusion, data level; Fusion, image level

Definition

Sensor level fusion combines raw biometric information that can account for inter-class and intra-class variability and facilitate decision making based on the fused raw information. A typical sensor level fusion algorithm first integrates raw biometric data either obtained from different viewpoints (for example, mosaicing several fingerprint impressions) or obtained from different sensors (for example, multimodal biometric images). The integrated data is then processed and discriminatory biometric features are extracted for matching. This level of fusion can be operated in both verification and ► identification modes. Few examples of sensor level fusion are: fingerprint mosaicing, multispectral face image fusion, and multimodal biometric image fusion.

Introduction

The concept of biometric information fusion is motivated from classical multi-classifier systems that combine information from different sources and represent using a single entity. Performance driven systems that use multiple biometric characteristics are known in multibiometric system [1]. These systems have several advantages over unimodal biometric systems such as tolerance to noise and malfunction, universality, and improved accuracy. Multibiometric systems are broadly classified into five levels of fusion.

1. *Sensor level fusion.* Raw data obtained directly from the sensors are fused without any feature extraction and represented as a single unit. This level of fusion is also known as data level fusion or image level fusion (for image based biometrics).
2. *Feature level fusion.* Data obtained from different sensors are first subjected to feature extraction algorithms and the feature sets are combined to generate a new feature vector which is subsequently used for recognition.
3. *Match score level fusion.* Features extracted from individual biometric modalities are first matched to compute the corresponding match scores. Match scores obtained from different biometric systems are then combined to generate a fused match score.
4. *Decision level fusion.* Decisions of individual biometric classifiers are fused to compute a combined decision. This level of fusion is also known as abstract level fusion because it is used when there is access to only decisions from individual classifier's.
5. *Rank level fusion.* With identification systems, rank level fusion involves combining identification ranks obtained from multiple unimodal biometrics. The output of rank level fusion is a consolidated rank that is used for final decision.

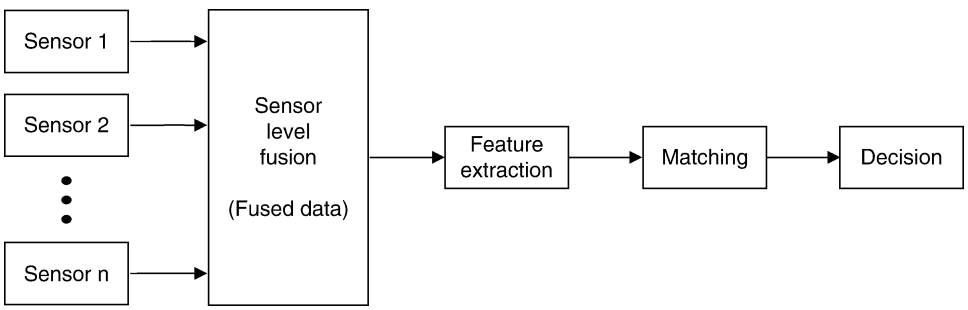
This article focuses on sensor level fusion and provides a comprehensive overview of the methodologies involved. In this level of fusion, first the raw data obtained from the sensors are combined to generate a fused data. An application oriented feature extraction algorithm is then used to compute the features from the fused data and matching is performed. Figure 1 illustrates the basic concept of sensor level fusion.

Sensor level fusion can be broadly classified into three categories: (1) single sensor multi-samples, (2) multi-sensor, and (3) multimodal. This article is organized to accentuate various algorithms proposed in each fusion category.

Sensor Level Fusion: Single Sensor Multi-Samples

In these systems, multiple samples of a single biometric modality are acquired using a single sensor and the information is combined to account for variations that can occur in a biometric modality. For example, as shown in Fig. 2, different profiles of a face image can be combined to obtain a fused representation of face image that can address the challenges due to pose variations [2]. In this category of sensor fusion, image ▶ mosaicing techniques are used for integrating information obtained from several impressions or view points, to augment the biometric content and to enhance the verification/identification performance. Singh et al. [2] describe the concept of mosaicing in biometrics as an exercise in information fusion when multiple images of a subject's biometric information are fused into a single entity in the image domain itself. Therefore, this could be viewed as fusion at the sensor level.

Mosaicing was first introduced in biometrics by Ratha et al. [3]. A rolled fingerprint image is generated from several partial fingerprint images using segmentation and blending algorithms assuming that the partial fingerprints are spatially registered. The performance of this fingerprint mosaicing algorithm is evaluated using different blending algorithms. The mosaicing algorithm generates rolled fingerprint image that is very close to the ground truth and improves the minutiae count that is useful for recognition. Further, Jain and Ross [4] proposed the use of iterative closest point algorithm to seamlessly register the ridges of two fingerprint images and to generate a composite fingerprint image. Recently, Ross et al. [5] employed thin-plate splines (TPS) to model the non-linear deformation in fingerprint images and integrate it



Fusion, Sensor-Level. Figure 1 Basic concept of sensor level fusion.

Fusion, Sensor-Level. Figure 2 Combining profile and frontal face images using mosaicing technique. (**a**) Profile and frontal face images and (**b**) Mosaiced face image.

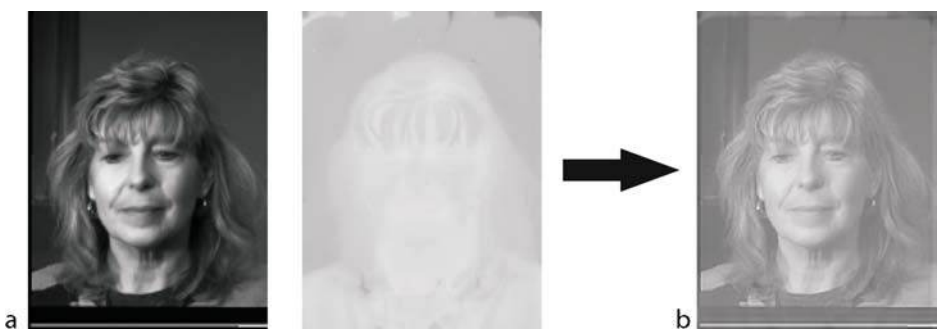
in the mosaicing process. This algorithm first aligns two fingerprint images using coarse alignment (affine model) followed by TPS based fine alignment. Once the fingerprint images are registered, a mosaiced fingerprint image is obtained by applying simple pixel averaging based blending method.

Mosaicing has also been applied to face biometrics. Yang et al. [6] describe an algorithm to create panoramic face mosaics. The acquisition system consists of five cameras that simultaneously obtains five different views of a subject's face. Based on the manually marked control points, the algorithm uses a series of linear transformations and smoothing operations on component images to generate a face mosaic. Unlike fingerprint mosaicing, face mosaicing requires specific feature extraction algorithm. Two different schemes to represent the panoramic image were proposed: one in the spatial domain and another in the frequency domain. Experimental evaluation on a database of 12 individuals shows that the face mosaicing algorithm improves identification accuracy in both the spatial and frequency domains. In [7], Liu and Chen describe a face mosaicing algorithm in which the human head is approximated with a 3D ellipsoidal model. The face, at a certain pose, is viewed as a 2D projection of this 3D ellipsoid. All 2D face images of a subject are projected onto this ellipsoid via geometrical mapping to form a texture map which is represented by an array of local patches. Matching is accomplished by adopting a probabilistic model to compute the distance of patches from an input face image. An identification accuracy of 90% on different databases has been reported. In [2], Singh et al. proposed a face mosaicing algorithm that can perform mosaicing in visible spectrum domain as well as in short wave infrared domain. The algorithm first registers the component face images using two stage registration algorithm and then a face mosaic is generated using multi-resolution splines based blending algorithm. Facial features are encoded using a generic feedforward hierarchical model-based feature extraction algorithm that extracts local facial features using the fundamentals of a biological visual system. Experiments conducted on three different face databases indicate that the proposed face mosaicing algorithm offers significant benefits by accounting for pose variations that are commonly observed in face images. Moreover, the mosaicing algorithm requires less time for matching compared to the score level fusion and also reduces the memory requirements.

Sensor Level Fusion: Multi-Sensors

In this category of sensor level fusion, multiple samples of a single biometric modality are obtained using multiple sensors and the information is combined such that the fused multi-sensor information improves the recognition performance. In general, the information obtained from multiple sensors are complementary to each other and can account for the intra-class variability. For example, as shown in Fig. 3, multi-spectral face images obtained using visible spectrum and infrared sensors can be fused to minimize the intra-class variations due to illumination and expression.

Multi-spectral face image fusion is the classical model for this level of fusion. Face recognition algorithms generally use visible spectrum images for recognition because the reflectance property yields a



Fusion, Sensor-Level. **Figure 3** Multi-spectral face image fusion. (**a**) Visible and infrared spectrum images and (**b**) Fused image.

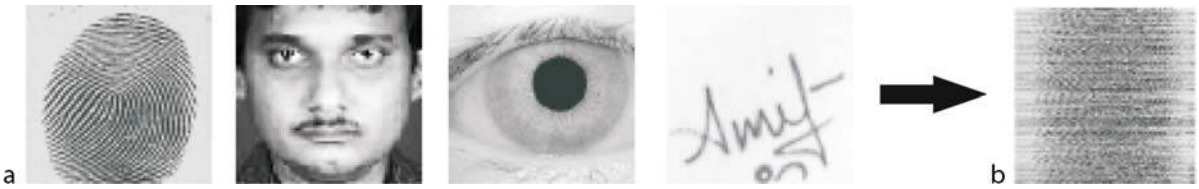
clear representation of facial features to differentiate between two individuals. However, visible spectrum images also possess several other properties which affect the performance of recognition algorithms. For example, changes in lighting affect the representation of visible spectrum images and can influence feature extraction. Other variations in facial appearance such as hairs, wrinkles, and expression are also evident in visible spectrum images and these variations increase the false rejection rate of face recognition algorithms. To address the challenges posed by visible spectrum images, researchers have used infrared images for face recognition [8]. Among all infrared spectrum images, long wave infrared (LWIR) images possess several properties that are complementary to visible images. Visible spectrum captures the electromagnetic energy in the range 0.4–0.7 μm, whereas long wave infrared or thermal images are captured in the range of 8–12 μm. Thermal images represent the heat pattern of the object and are invariant to illumination and expression. Face images captured in long wave infrared spectrum have less intra-class variation and help to reduce the false rejection rate of recognition algorithms. These properties of long wave infrared and visible images can be combined to improve the performance of face recognition algorithms.

In literature, researchers have proposed several multi-spectral face image fusion algorithms [8]. Bebis et al. [9] proposed an image fusion algorithm in wavelet domain using genetic algorithm. In this algorithm, multi-spectral face images are first transformed into wavelet domain and a multiresolution representation is obtained. Then, a genetic algorithm is used to select the most appropriate wavelet coefficients at pixel level. Finally, inverse ▶ wavelet transform is applied to generate a fused face image and Eigenface based algorithm is used for feature extraction and matching. The genetic fusion algorithm suffers from making a good choice of fitness function. Kong et al. [10] proposed a wavelet based multi-spectral face image fusion algorithm in which the visible and infrared spectrum images are first registered using affine transformation. An empirical weighting scheme is then applied on the registered multi-spectral face images in wavelet domain to obtain the composite face image. Although the algorithm is straightforward, the generic empirical weighting scheme is not sufficient to address the inter-class and intra-class variability in face images. Recently, Singh et al. [11] proposed a 2ν-granular support vector machine (2ν-GSVM) based multi-spectral face image fusion algorithm. This algorithm first registers multi-spectral face images using mutual information based registration algorithm. Then, a 2ν-GSVM learning scheme is invoked in wavelet domain to learn the properties of the multi-spectral face images at different resolution and granularity levels, determine optimal information and combine them to generate a fused image. Finally, texture features are extracted from the fused image for recognition. Experimental results show that 2ν-GSVM based fusion algorithm can address the challenges due to illumination, expression, and occlusion variations. This algorithm provides improved verification accuracy ($>94\%$) compared to other image fusion schemes.

Another example of sensor level fusion with multiple sensors is fusing 2D and 3D facial information. Lu et al. [12] describe a semi-automatic sensor level fusion algorithm that integrates range and texture features for improved face recognition performance. Combining 3D shape information with registered 2D texture information using iterative closest point algorithm improves the face identification performance. The authors report that the algorithm is robust to arbitrary view, lighting, and facial appearance. However, the algorithm is computationally expensive and suffers due to non-rigid variations.

Sensor Level Fusion: Multimodal

In most of the multimodal biometric systems, such as bimodal system with face and fingerprint, fusion is performed at match score level or decision level. Very limited research is undertaken to perform sensor level fusion in a multimodal system. In this category of sensor level fusion, multimodal biometric images are fused to address issues such as universality, memory storage, small sample size recognition, and recognition performance. Further, an efficient sensor level fusion algorithm has advantages due to the availability of fused raw information from where the representative composite biometric features can be extracted and used for matching. The main challenge lies in developing fusion algorithm that can account for inter-class and intra-class variability in multimodal biometric images. An example of multimodal biometric image fusion is shown in Fig. 4.



Fusion, Sensor-Level. **Figure 4** Multimodal image fusion. (**a**) Image pertaining to different biometric modalities and (**b**) Fused and scrambled image.

Jing et al. [13] propose a sensor level fusion algorithm that generates a composite image from face and palmprint biometrics. Circular Gabor filters are first applied on face and palm print images to generate 32 filtered responses of each biometric data. These filtered responses are concatenated to generate a fused image. A pixel normalization scheme is then used to minimize variations due to imaging conditions. Finally, kernel discriminative common vectors are extracted from the fused image and radial basis function based neural network is used for classification. The fusion algorithm improves the recognition performance and is an effective solution for the small sample size recognition problem. Noore et al. [14] proposed discrete wavelet transformation based image fusion algorithm that generates a composite image by combining multimodal biometric images. The algorithm starts with transforming biometric images into wavelet domain and generating composite image by amalgamating the wavelet coefficients. The composite image is then scrambled using a secret encoding key generated with Fibonacci transforms. The algorithm not only improves the recognition performance but also reduces the memory requirements and provides resilience to common image processing attacks such as smoothing, cropping, JPEG 2000 compression, and filtering.

Future Research Directions

As discussed in previous sections, sensor level fusion has several advantages. However, compared to other levels of fusion, this level of fusion is less explored and requires further research to address the limitations of current research. First and foremost is to further improve the recognition accuracy. Researchers have shown that for certain applications, sensor level fusion algorithms do not provide better results compared to match score level fusion algorithms [5, 11]. This is mainly because existing algorithms do not effectively reconcile the information that is useful for recognition. We believe that existing sensor level fusion algorithms fail in some cases because during information fusion it is possible that redundant and less discriminatory features become predominant. Furthermore, there is a lack of generalized sensor level fusion algorithms that can be used for different biometric scenarios or applications. For instance, genetic algorithm based multi-spectral image fusion algorithm can not be directly used for multimodal image fusion. Additional research is required to design an effective and generalized sensor level fusion algorithm which can be applied to different biometric modalities. Every sensor level fusion algorithm requires specific feature extraction algorithm that can effectively extract discriminatory biometric information from the composite image or data. This requirement is not mandatory with a generalized sensor level fusion algorithm. Therefore, a generalized algorithm can be easily incorporated in commercial systems and can conform to data fusion standards.

Another important research issue is to unify the sensor level fusion in a ▶ unification framework that reconciles multiple fusion algorithms. Originally proposed by Vatsa et al. [15], a biometric unification framework combines multiple fusion algorithms by dynamically selecting the most appropriate fusion algorithm depending on the input evidences such as quality and other priors. Currently, the unification framework includes only the match score fusion algorithms. However, with proper modifications, the unification framework can be expanded to include multi-level fusion algorithms that can address the operational needs of biometric systems and provide better recognition performance.

Related Entries

▶ Data Fusion
▶ Face Recognition

► Identification
► Verification

References

1. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. 1st edn. Springer, New York (2006)
2. Singh, R., Vatsa, M., Ross, A., Noore, A.: A mosaicing scheme for pose-invariant face recognition. IEEE Trans. Syst. Man Cybern. Part B **37**(5), 1212–1225 (2007)
3. Ratha, N.K., Connell, J.H., Bolle, R.M.: Image mosaicing for rolled fingerprint construction. In: Proceedings of International Conference on Pattern Recognition, pp. 1651–1653 (1998)
4. Jain, A., Ross, A.: Fingerprint mosaicking. In: Proceedings of International Conference on Acoustics, Speech, and Signal Processing, pp. 4064–4067 (2002)
5. Ross, A., Shah, S., Shah, J.: Image versus feature mosaicing: a case study in fingerprints. In: Proceedings of SPIE Conference on Biometric Technology for Human Identification III, pp. 620,208-1–620,208-12 (2006)
6. Yang, F., Paindavoine, M., Abdi, H., Monopoly, A.: Development of a fast panoramic face mosaicing and recognition system. Opt. Eng. **44**(8), 087 005/1–087 005/10 (2005)
7. Lu, X., Jain, A.K.: Pose-robust face recognition using geometry assisted probabilistic modeling. In: Proceedings of International Conference on Computer Vision and Pattern Recognition, pp. 502–509 (2005)
8. Kong, S., Heo, J., Abidi, B., Paik, J., M., A.: Recent advances in visual and infrared face recognition - a review. Comput. Vision Image Understand. **97**(1), 103–135 (2005)
9. Bebis, G., Gyaourova, A., Singh, S., Pavlidis, I.: Face recognition by fusing thermal infrared and visible imagery. Image Vision Comput. **24**(7), 727–742 (2006)
10. Kong, S., Heo, J., Boughorbel, F., Zheng, Y., Abidi, B., Koschan, A., Yi, M., M., A.: Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition. Int. J. Comput. Vision **71**(2), 215–233 (2007)
11. Singh, R., Vatsa, M., Noore, A.: Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition. Pattern Recognit. **41**(3), 880–893 (2008)
12. Lu, X., Jain, A.K.: Integrating range and texture information for 3D face recognition. In: Proceedings of Workshop on Applications of Computer Vision, pp. 156–163 (2005)
13. Jing, X.Y., Yao, Y.F., Zhang, D., Yang, J.Y., Li, M.: Face and palmprint pixel level fusion and Kernel DCV-RBF classifier for small sample biometric recognition. Pattern Recognit. **40**(11), 3209–3224 (2007)
14. Noore, A., Singh, R., Vatsa, M.: Robust memory efficient data level information fusion of multi-modal biometric images. Inf. Fusion **8**(4), 337–346 (2007)
15. Vatsa, M., Singh, R., Noore, A.: Unification of evidence theoretic fusion algorithms: A case study in level-2 and level-3 fingerprint features. In: Proceedings of IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp. 1–6 (2007)

Fusion, User-Specific

Julian Fierrez, Javier Ortega-Garcia
Biometric Recognition Group – ATVS, Escuela Politecnica Superior, Universidad Autonoma de Madrid, Campus de Cantoblanco, Madrid, Spain

Synonyms

Adapted fusion; Local fusion; Target-dependent fusion; User-dependent fusion

Definition

User-specific fusion in the framework of biometrics, initially devised for score fusion in the verification mode, refers to techniques used for information fusion in which there is a specific fusion function for each user enrolled in the system. These fusion functions are retrieved and used for information integration in the same way the enrolled templates corresponding to the claimed identities are retrieved and used for matching.

User-specific fusion techniques find application in several biometric fusion scenarios, e.g., multi-modal fusion, where some subjects may be not adequate for recognition based on specific modalities (these evidences can be ignored or given less importance in the information fusion step), or multi-algorithm fusion, where some subjects may be better recognized based on particular algorithms (their fusion functions can be adapted to give more importance to those algorithms).

The biggest challenge for effective user-specific fusion is the need for user-specific training data, which is usually very scarce. Recent user-specific fusion techniques exploit the usually scarce training data by considering also for training the information provided by background users. These new techniques are known as adapted user-specific fusion.

System Model

The following nomenclature is used throughout the essay. Given a multi-biometric verification system consisting of a number of uni-modal systems, each one computes a similarity score between an input biometric pattern and the enrolled pattern or model of the

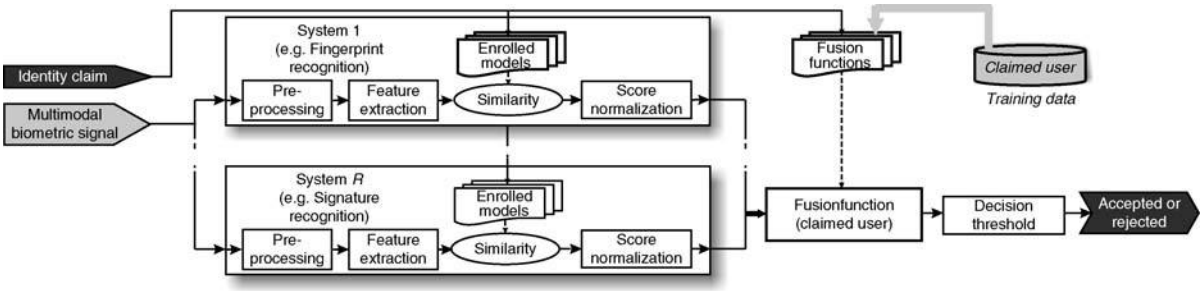
given claimant. The similarity scores are then normalized to a given score range. Let the normalized similarity scores provided by the different uni-modal systems be combined into a multi-modal score. The design of a fusion scheme consists in the definition of a function which maps a multi-modal score to a fused real value, so as to maximize the separability of client and impostor fused score distributions. This function may be fixed or trained (see the entry in this encyclopedia on Multi-biometrics) by using a set of training scores (scores known to be genuine or impostor).

The aim in user-specific fusion is to obtain the best score fusion function for a particular user, resulting in the system model shown in Fig. 1.

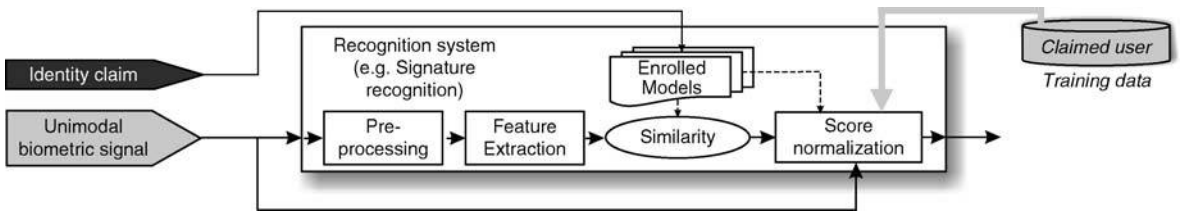
User-Specific Multi-Biometrics

User-specific multi-biometric verification can be achieved not only by making the fusion functions user-specific as shown in Fig. 1, but also other processing modules, such as the score normalization and the decision processing blocks. In the first case, each individual system will be used as indicated in Fig. 2, in the latter case the overall system diagram will be as indicated in Fig. 3.

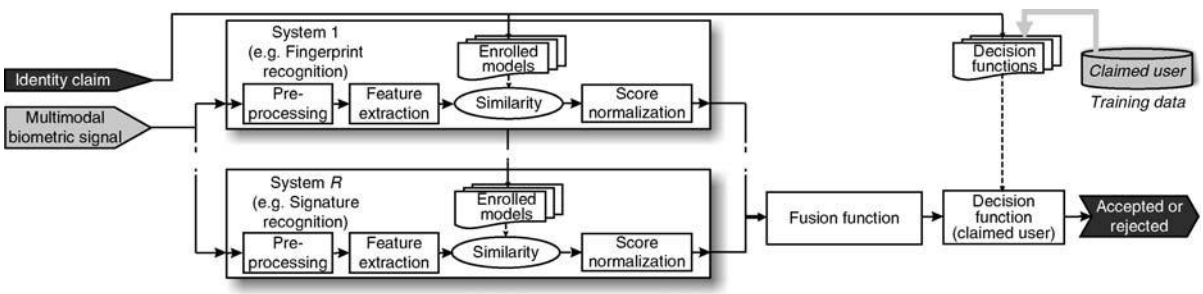
On one hand, user-specific score normalization has been traditionally studied for individual behavioral biometric modalities in which there are large variations between users (such as speech [1] or signature [2]),



Fusion, User-Specific. **Figure 1** System model of multi-biometric verification with user-specific score fusion.



Fusion, User-Specific. **Figure 2** System model of biometric verification with user-specific score normalization.



Fusion, User-Specific. **Figure 3** System model of multi-biometric verification with user-specific decision functions.

where their application is very effective to compensate the problems related to the heterogeneity between users. When user-specific score normalization is used in one of the systems being combined in a multi-biometric setup, the resulting approach can be seen as integrating the multi-biometric data in a user-specific way [3]. Despite the success of user-specific score normalization in individual modalities, and the success of fusion techniques, few efforts have been reported in the literature studying the combined use of both techniques to make the most out of the usually scarce user-specific training data.

On the other hand, the use of user-specific decisions in multi-biometrics has been typically studied in combination with user-specific score fusion. In this case, it has been demonstrated that it is better to use the available training data for computing user-specific fusion functions instead of user-specific decision schemes [4].

User-Specific Fusion

The idea of exploiting user-specific parameters at the score level in multi-modal biometrics was introduced, to the best of our knowledge, by [5]. In that work, user-independent weighted linear combination of similarity scores was demonstrated to be improved by using either user-specific weights or user-specific decision thresholds, both computed by exhaustive search on the testing data. The idea of user-specific fusion parameters was also explored by [6]. Other attempts to personalize multi-modal biometrics include the use of the claimed identity index as a feature for a global trained fusion scheme based on neural networks [7], computing user-specific weights using lambness metrics [8], and using personalized Fisher ratios [9].

The existing score fusion approaches can be classified as global or local depending first on the fusion function (i.e., user-independent or user-specific fusion strategies) and secondly on the decision making process (i.e., user-independent or user-specific decision thresholds), resulting in [10]: global-learning-global-decision (GG), local-learning-global-decision (LG), and similarly GL and LL. Some example works on user-specific multi-biometrics using this classification are: LG [4, 5, 6, 7, 8, 10, 11], GL [4, 5, 10], and LL [4, 10].

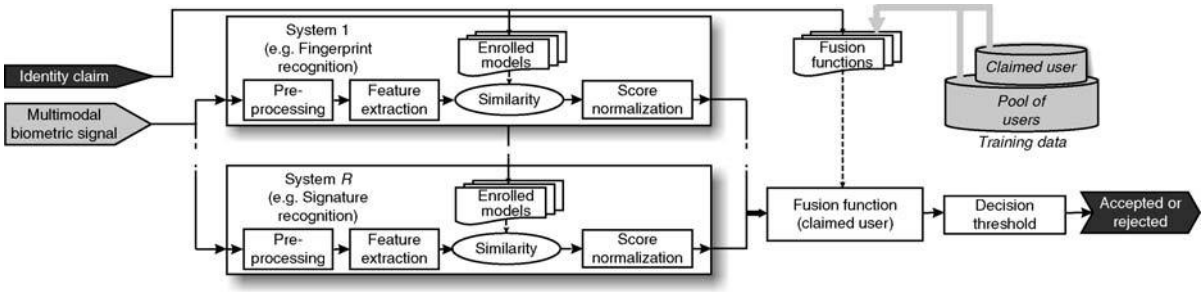
User-specific score fusion is confronted with a great challenge: the scarcity of user-specific training scores. For overcoming this challenge, the simultaneous use of user-specific and background information has been proposed for training the user-specific fusion functions, in what has been called *adapted user-specific fusion*. This approach can be seen as a particular case of a more general type of approaches, referred to as ▶ adapted fusion [11]. In these approaches, a baseline fusion function is first constructed based on some general knowledge of the problem at hand, and then adjusted during the operation of the system. The adaptation can be based on ancillary information such as: the user being claimed (adapted user-specific fusion), quality measures of the input biometrics (quality-based fusion [12], see related entry in this encyclopedia), or other kind of environmental information affecting the various information channels being fused.

Adapted User-Specific Fusion

Adapted methods in the context of user-specific fusion refer to the use of both global and local information for learning the fusion functions.

The idea of adapted learning is based on the fact that the amount of available training data in localized learning is usually not sufficient and representative enough to guarantee good parameter estimation and generalization capabilities. To cope with this lack of robustness derived from partial knowledge of the problem, one can exploit the information provided by background global data. In general, the relative balance between the background information (pool of users) and the local data (specific user) is performed as a tradeoff between both kinds of information.

The system model of adapted user-specific score fusion is shown in Fig. 4, where we can see that the fusion function of a given user is trained with two sets of training data, both including both genuine and impostor matching scores. The first training set consists of scores corresponding to the user being claimed. The second set consists of scores corresponding to a pool of background users different to the user being claimed. By considering these two sets simultaneously, the resulting adapted user-specific fusion schemes outperform the traditional user-independent fusion (also known as ▶ global fusion, in which only the pool of users is used for training), and the traditional user-specific fusion depicted in Fig. 1 (also known as ▶ local fusion, in which only data from the claimed



Fusion, User-pecific. **Figure 4** System model of multi-modal biometric authentication with adapted user-specific score fusion.

user is used for training). This affirmation has been demonstrated experimentally in various scenarios, such as multi-algorithm speaker verification [3, 13], and multi-modal verification combining on-line signature and fingerprint traits [4, 11].

Related Entries

- ▶ Fusion, Quality-Based
- ▶ Multi-Algorithm Systems
- ▶ MultiBiometrics
- ▶ Multi-Modal Systems

References

1. Doddington, G., Liggett, W., Martin, A., Przybocki, M., Reynolds, D.: Sheeps, goats, lambs and wolves: a statistical analysis of speaker performance in the NIST 1998 speaker recognition evaluation. In: Proceedings of International Conference on Speech and Language Processing, ICSLP (1998)
2. Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Target dependent score normalization techniques and their application to signature verification. IEEE Trans. Syst. Man Cybernet., Part C **35**, 418–425 (2005)
3. Poh, N., Kittler, J.: Incorporating model-specific score distribution in speaker verification systems. IEEE Trans. Audio Speech Lang. Process. **16**, 594–606 (2008)
4. Fierrez-Aguilar, J., Garcia-Romero, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Adapted user-dependent multimodal biometric authentication exploiting general information. Pattern Recognit. Lett. **26**, 2628–2639 (2005)
5. Jain, A.K., Ross, A.: Learning user-specific parameters in a multibiometric system. In: Proceedings of IEEE International Conference on Image Processing, ICIP, vol. 1, pp. 57–60 (2002)
6. Wang, Y., Wang, Y., Tan, T.: Combining fingerprint and voice biometrics for identity verification: An experimental comparison. In: Zhang, D., Jain, A.K. (eds.) Proceedings of International Conference on Biometric Authentication, ICBA, Springer LNCS-3072, pp. 663–670 (2004)
7. Kumar, A., Zhang, D.: Integrating palmprint with face for user authentication. In: Proceedings of Workshop on Multimodal User Authentication, MMUA, pp. 107–112 (2003)
8. Snelick, R., Uludag, U., Mink, A., Indovina, M., Jain, A.K.: Large scale evaluation of multimodal biometric authentication using state-of-the-art systems. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 450–455 (2005)
9. Poh, N., Bengio, S.: An investigation of f-ratio client-dependent normalisation on biometric authentication tasks. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, vol. 1, pp. 721–724 (2005)
10. Toh, K.A., Jiang, X., Yau, W.Y.: Exploiting local and global decisions for multimodal biometrics verification. IEEE Trans. Signal Process. **52**, 3059–3072 (2004)
11. Fierrez-Aguilar, J., Garcia-Romero, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Bayesian adaptation for user-dependent multimodal biometric authentication. Pattern Recognit. **38**, 1317–1319 (2005)
12. Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., Bigun, J.: Discriminative multimodal biometric authentication based on quality measures. Pattern Recognit. **38**, 777–779 (2005)
13. Fierrez-Aguilar, J., Garcia-Romero, D., Ortega-Garcia, J., Gonzalez-Rodrigez, J.: Speaker verification using adapted user-dependent multilevel fusion. In: Oza, N.C., Polikar, R., Kittler, J., Roli, F. (eds.) Proceedings of International Workshop on Multiple Classifier Systems, MCS, Springer LNCS-3541, pp. 356–365 (2005)

Fusion, Wavelet-Based

Wavelet-based fusion has been widely used in literature. The wavelet transform is a data analysis tool that provides a multi-resolution decomposition of an image. Wavelet-based pixel-level data fusion is used on two or more sets of probe images. Given two

registered images I_1 and I_2 of the same object from two sets of probe images (two different spectral bands in this case), a two-dimensional discrete wavelet decomposition is performed on I_1 and I_2 to obtain the wavelet approximation coefficients (a_1, a_2) and detail coefficients (d_1, d_2). The wavelet approximation and detail coefficients of the fused image, a_f and d_f, are then calculated as follows:

$$a_f = W_{a_1} \times a_1 + W_{a_2} \times a_2 \quad \text{and}$$
$$d_f = W_{d_1} \times d_1 + W_{d_2} \times d_2,$$

where $W_{a_1}, W_{a_2}, W_{d_1}$, and W_{d_2} are weights determined either empirically or according to some selected rule. The two-dimensional discrete wavelet inverse transform is then performed to obtain the fused image.

► Multispectral and Hyperspectral Biometrics

Fuzzy Extractor

► Encryption, Biometric
► Fingerprints Hashing

Fuzzy Vault

Fuzzy vault is where a secret key is hidden behind some biometric data which are fuzzy and noisy by nature.

► Fingerprints Hashing

G

Gabor Jets

Gabor jets are a set of filters that are used to extract the local frequency information from the face images. These filters are generally linear filter with impulse responses defined by a harmonic function and a Gaussian function. The Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function.

► Face Recognition, Component-Based

Gabor Transform

A complete representation of a signal or image in terms of coefficients on Gabor wavelets, such that the original data can be reconstructed exactly by combining together those wavelets using their computed coefficients. A complication is that the necessary coefficients cannot be obtained simply by operations of filtering or by the inner product projections of the data with the wavelets, since they do not constitute an orthogonal basis. More complex methods are required (biorthogonal bases; relaxation networks) to obtain the needed expansion coefficients from projection coefficients. Once obtained, a Gabor Transform is a powerful tool for signal or image encoding, analysis, and compression.

► Iris Encoding and Recognition using Gabor Wavelets

Gabor Wavelets

Complex exponentials (Fourier components) multiplied by Gaussian envelopes. Although they fail to satisfy some parts of the stricter mathematical definitions of wavelets, such as orthogonality and compact support, these elementary functions can constitute a powerful basis for signal or image encoding, representation, compression, and analysis. They are increasingly used today in computer vision and in pattern recognition, particularly in biometrics, where they are the basis of iris recognition and have also been used for several other biometric modalities. Among their advantages (besides forming a complete basis for signal or image encoding) are; their optimality under the Heisenberg Uncertainty Principle for simultaneous resolution in time/space and in frequency; their closed analytical form; their self-Fourier property and closure under convolution and multiplication; and their neurobiological basis in the receptive field profiles of neurons in the mammalian visual cortex. Their chief disadvantage is that they are not mutually orthogonal, and so the projection coefficients obtained by computing their inner product with an image are not the same as the expansion coefficients that would be needed to reconstruct the same image exactly from them.

► Face Recognition, Component Based
► Iris Encoding and Recognition using Gabor Wavelets
► Local Image Features
► Local Image Filters

© 2009 Springer Science+Business Media, LLC

Gait

The manner of a person's movement, specifically during walking is called gait. The human gait cycle consists of two main phases: during stance phase, the foot is on the ground, and during the swing phase, the leg is swinging forward in preparation for the next ground contact.

► Gait, Forensic Evidence of

Gait Analysis

► Gait, Forensic Evidence of

Gait Biometrics, Overview

Rama Chellappa, Ashok Veeraraghavan
Narayanan Ramanathan
University of Maryland, College Park, MD, USA

Synonym

Gait recognition

Definition

Gait is defined as the style or manner of walking. Studies in physchophysics suggest that people can identify familiar individuals using just their gait. This has led to a number of automated vision based algorithms that use gait as a biometric. Such a system usually consists of a video camera capturing images of a person walking within its field of view. Appropriate features such as joint angles or silhouettes are extracted from this video and are then used to compare with the stored gait signatures of known individuals. As with any other biometric system, the system can operate in both the identification and the verification mode. Gait as a biometric has several advantages compared to traditional biometrics such as fingerprint in that gait is non-intrusive, does not require cooperation from the individual, and can function at moderate distances from the subject.

Introduction

The study of human gait has gathered pace in recent years driven primarily by its potential as a biometric. Gait-based person authentication has several significant advantages compared to traditional biometrics such as fingerprint or iris. Firstly, gait based biometric systems do not require the individuals to be cooperative since the input of these systems is the video feed captured by passive cameras. Secondly, gait is a non-intrusive biometric – it does not require the individuals to wear any special equipment in order to be recognized. Thirdly, gait based biometric systems have an extended range compared to traditional biometrics – they can operate reliably even when the subjects are tens of meters away from the camera. Finally, such a system harnesses the potential of thousands of surveillance video cameras installed in public locations into a biometric authentication system.

Operation of a Gait Based Biometric System

The sensor for a gait-based biometric system is a video camera capturing videos of human subjects walking within its field-of view. The raw sensor video is then processed to extract relevant features which can then be used for recognition. If the acquisition conditions are expected to be controlled and favorable, then the quality of the video will enable the extraction of features such as joint angles from the individual video frames. In more typical uncontrolled settings, the features extracted could either be background subtracted binary images, silhouettes, shapes or width vectors – all examples of features capturing the extent of the human body to differing amounts of detail. During the training phase, several such sequences of each individual in the gallery are collected and the appropriate features are then stored in the database. During the test phase, each test sequence is compared with the training sequences available in the database and the similarity is used to perform person authentication.

Challenges for Gait Based Biometrics Systems

The discriminative information in gait is present in both the shape of the individual and also in the manner of his/her gait. This means that gait based biometric systems must be able to model gait as a time series of features or as a dynamical model in order to perform accurate recognition. Static template based methods which have been used for most other biometric systems need to be adapted to a temporal sequence in order to achieve robust performance. In this regard, another challenge is time alignment of two sequences so that critical events during gait like "mid-stance", "toe-off " etc. are time aligned accurately so that recognition performance is not affected by inaccurate time alignment between postures that occur during gait. Since, gait based person identification often occurs without any particular viewpoint, view-invariance of the feature extracted from the video is another important challenge. This will ensure that recognition performance is robust to changes in the viewpoint of the camera. In scenarios with moderate amounts of acquisition control, one can set up multiple video cameras so as to ensure that the best possible viewpoint which happens to be the fronto-parallel gait is captured on atleast one of the cameras. Another challenge for automated gait-based biometrics is that of changing illumination conditions in the scene. In order to be robust to changing illumination conditions, background subtraction is typically performed on the raw videos before the video data is used in a recognition algorithm. Finally, another important challenge is the variability in the clothing, shoe type and the surface on which the individuals walk. Obviously, the clothing of the subject especially their type of footwear has significant impact on the gait features observed and it is important to bear this in mind while developing gait-based biometric systems.

Features for Gait Based Biometrics

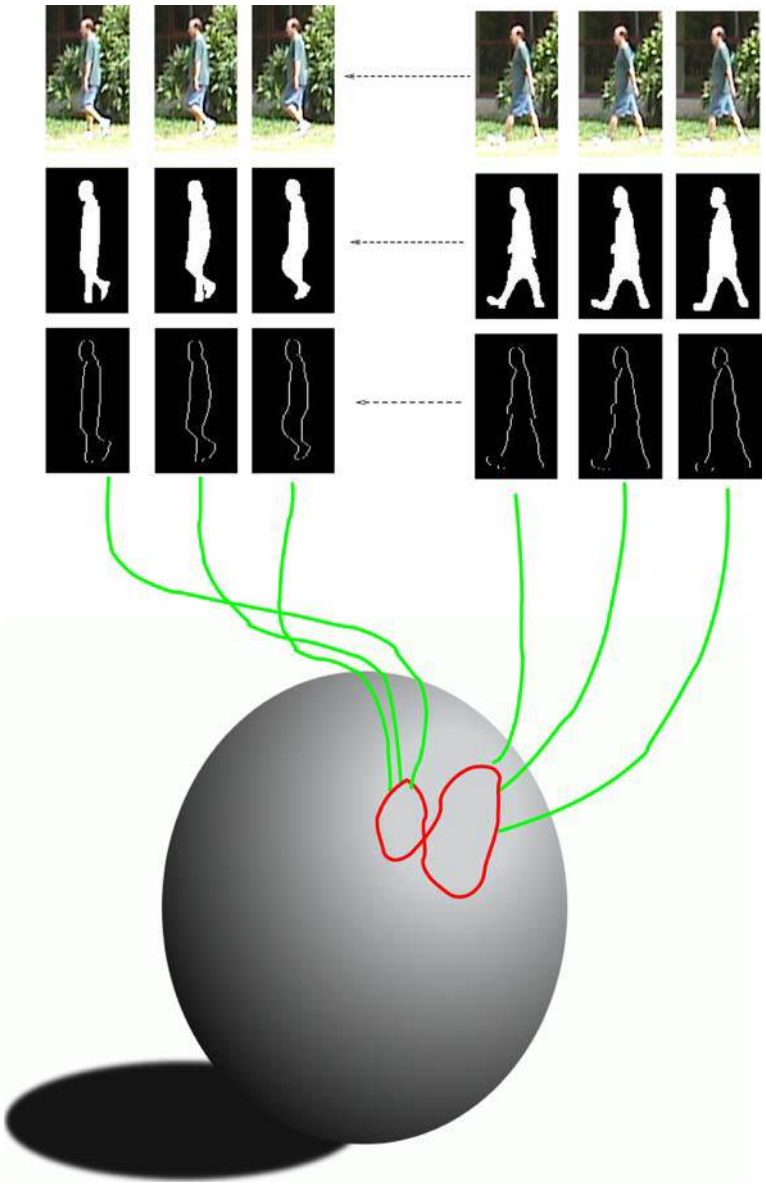
Silhouette: In most gait-based biometric systems the cameras can be assumed to be static during the short duration of time that they capture the gait of a single individual for verification. This allows simple background models to be built for each of these cameras. Background subtraction then identifies the set of all pixels in the image that belong to the moving individual. Figure 1 shows a sequence of color images captured by a video camera as a person walks through its field of view. Shown below are the binary background subtracted images in which all pixels belonging to the individual are white, while the background is black. This binary image is then scaled to a uniform size so that the feature extracted is independent of the distance of the camera from the subject. Several algorithms for gait based person identification use this binary silhouette as a feature [1–5].

Shape: "Shape is all the geometric information that remains when location, scale, and rotational effects are filtered out from the object"[6]. Kendall's statistical shape is a sparse descriptor of the shape that describes the shape configuration of k landmark points in an m-dimensional space as a $k \times m$ matrix containing the coordinates of the landmarks. Image space is 2-dimensional and therefore it is convenient to describe the shape vector as a k dimensional complex vector. First, a binarized silhouette denoting the extent of the object in an image is obtained. A shape feature is then extracted from this binarized silhouette. This feature vector must be invariant to translation and scaling since the object's identity should not depend on the distance of the object from the camera. So any feature vector that we obtain must be invariant to translation and scale. This yields the pre-shape of the object in each frame. Pre-shape is the geometric information that remains when location and scale effects are filtered out. Let the configuration of a set of k landmark points be given by a k-dimensional complex vector containing the positions of landmarks. Let us denote this configuration as X. Centered pre-shape is obtained by subtracting the mean from the configuration and then scaling to norm one. The centered pre-shape is given by

$$Z_c = \frac{CX}{\| CX \|}, \ \textit{where } C = I_k - \frac{1}{k} 1_k 1_k^T, \qquad (1)$$

where I_k is a k × k identity matrix and 1_k is a k dimensional vector of ones.

The advantage of using shape feature is that the differential geometric properties of the spherical manifold in which the shapes lie are very well understood and therefore, appropriate distance measures that can account for translational, rotational and scale invariances are well defined. For example, consider two complex configurations X and Y with corresponding preshapes α and β. The full Procrustes distance



Gait Biometrics, Overview. **Figure 1** Graphical illustration of a video sequence obtained during a walking cycle and the corresponding features – silhoeuette and shape. Courtesy [7].

between the configurations X and Y is defined as the Euclidean distance between the full Procrustes fit of α and β and is chosen so as to minimize

$$d(Y,X) = \| \beta - \alpha s e^{j\theta} - (a+jb)1_k \|, \qquad (2)$$

where s is a scale, θ is the rotation and $(a+jb)$ is the translation. The full Procrustes distance is the minimum Full Procrustes fit i.e.,

$$d_F(Y,X) = \inf_{s,\theta,a,b} d(Y,X). \qquad (3)$$

The extracted shape sequence is shown in the bottom row of Figure 1 with a graphical illustration of the spherical manifold in which shapes lie. Shape is a very popular feature for gait-based biometrics and several state of the art algorithms perform gait matching as a matching of a sequence of shapes [7–10].

Joint Angles: A very popular feature for gait analysis in the medical and the psychophysics community is the joint angles – i.e., the angles made at each of the limb joints such as the knee, elbow ankle, wrist

etc. There have been a few gait based biometrics algorithms that use joint angles as the feature for matching [11, 12]. The advantage of using joint angles as a feature is the fact that view-invariance is automatically achieved while using joint angles as a feature. Nevertheless, the essential problem with using joint angles is the fact that it is very challenging to robustly estimate them from uncontrolled monocular video sequences.

Algorithms for Matching

Most of the features described above have incorporated modest forms of view-invariance (atleast scale and translational invariance) as a part of the feature. Therefore the essential task of the algorithm for matching would be to model the dynamics of the feature during gait and use this to perform matching in a manner that is fairly insensitive to the speed of walking.

Dynamic Time Warping (DTW): Dynamic time warping is an algorithm for estimating the non-linear time synchronization between two sequences of features. The two sequences could be of differing lengths. Experiments indicate that the intra-personal variations in gait of a single individual can be better captured by non-linear warping rather than by linear warping [13]. The DTW algorithm which is based on dynamic programming computes the best non-linear time normalization of the test sequence in order to match the template sequence, by performing a search over the space of all allowed time normalizations. The space of all time normalizations allowed is cleverly constructed using certain temporal consistency constraints. Several gait-based biometrics algorithms have used the Dynamic time warping algorithm in order to time synchronize and match gait sequences [7, 8]. Recently, the DTW algorithm has also been extended so as to learn the warping constraints in a class-specific manner in order to improve discrimination between individuals [9].

Hidden Markov Model (HMM) The Hidden Markov Model (HMM) is a statistical state space model in which the observed shape sequence is modeled as outputs of a hidden states whose transitions are assumed to be Markovian. The model parameters of the HMM encode both the transition probabilities between the hidden states and the outputs of hidden states. The advantage of using a HMM is that there exists a wealth of literature on learning the parameters of the HMM and to perform inference using the HMM. Typically, the model parameters for each individual in the gallery is learnt and stored during the training phase. During the test phase, the probability of the observation sequence conditioned on the model parameters is maximized in order to perform recognition. The HMM [2, 3] and its many variants [14] have been successfully used for gait based person identification.

Autoregressive Moving Average Model (ARMA): Matching gait biometrics essentially is a problem of matching time-series data where the feature at each time instant is a silhouette or shape or joint angles. Therefore traditional time series modeling approaches such as the autoregressive model (AR) and the autoregressive moving average (ARMA) model have also been successfully used for gait based person identification. The model parameters of the ARMA model are learnt from the training sequences and stored. Given a test sequence, the model parameters for the test sequence are learnt and the distance between the model parameters is used in order to perform recognition [7].

Model Based Approaches

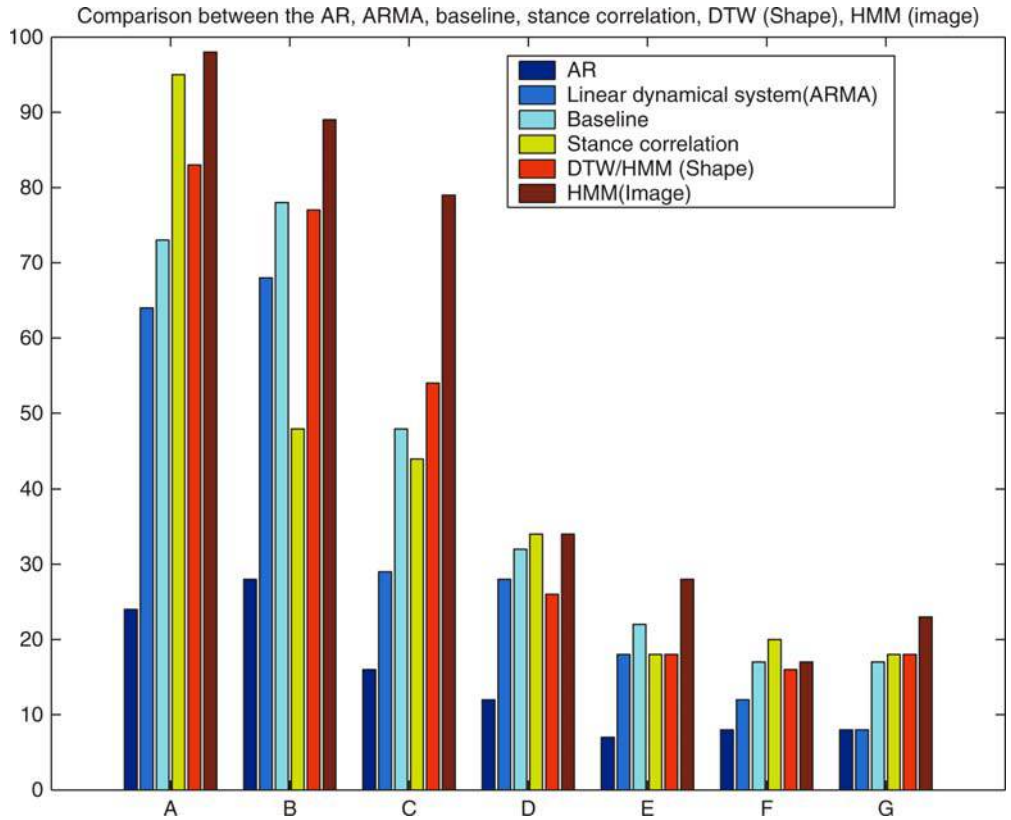
Typical feature based approaches first compute a sequence of features from each video and then match the sequence of features obtained in the test video to those stored in the gallery. Model-based approaches are different in the sense that they fit the sequence of features to a physical model of the human body and its inherent dynamics. For example, a model-based feature extraction process guided principally by biomechanical analysis for gait-based person identification is proposed [15]. The shape model for human subjects is composed of an ellipse to describe the head and the torso, quadrilaterals to describe the limbs and rectangles to describe the feet. Anatomical data is first used in order to derive shape and motion models that are consistent with normal human body proportions. Prototype gait motion models are then adapted to individuals using the specific characteristics of the extracted features. These individual specific shape and motion models are then used for gait recognition. A systematic analysis of the model-based approach also showed that cadence and static shape parameters of the human body account for most of the recognition performance.

Experiments on the USF Gait Data

In order to quantitatively test the performance and the viability of gait based biometrics a challenging gait database of 122 individuals was collected at the University of South Florida [4] as part of the DARPA Human Identification at a Distance (HID) program. The entire dataset containing over 1,200 videos was separated into 12 different experiments with varying levels of difficulty. The different challenge experiments amounted to varying different covariates during gait, like viewpoint, clothing, surface type, shoe type, and time etc. A bar plot of the recognition performance of various algorithms on the USF dataset (Experiments A-G) is shown in Figure 2. Experiments A,B and C correspond to changes in "view", "shoe type" and "view + shoe type" respectively without any change in the surface of walking, while challenege experiments D,E,F and G correspond to changes in the surface type from grass to concrete. The experiments indicate that changes in the surface type has significant impact on the recognition performance while view, shoe type affects recognition performance to a much lesser degree.

Summary

Gait is thus a novel biometric that provides significant operational advantages over several other biometrics such as face, fingerprint, iris etc. Unlike traditional biometrics like fingerprint, gait does not require the active cooperation of the subjects. Moreover, gait is a medium range biometric in the sense that acquisition distances can be as large as tens of meters. Moreover, in most operational scenarios, it is non-intrusive and does not require the subject to wear any special clothing. Preliminary experiments into gait as a biometric seem to indicate that the discriminative power of gait is not as strong as that of traditional biometrics such as fingerprints or iris. Therefore, several successful investigations for fusing the gait biometric with other



Gait Biometrics, Overview. Figure 2 Comparison of various algorithms on the USF gait database. (Courtesy [1]).

traditional biometrics in order to boost the identification performance have been performed and this seems to be an area of immense potential [16].

Related Entries

- ▶ Biometrics, Overview
- ▶ Covariates
- ▶ Multibiometrics
- ▶ Surveillance

References

1. Veres, G., Gordon, L., Carter, J., Nixon, M.: What image information is important in silhouette-based gait recognition? In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision Pattern Recognition (CVPR 2004), Washington, DC, 27 June 2004, vol. 2 (2004)
2. Kale, A., Sundaresan, A., Rajagopalan, A., Cuntoor, N., Roy-Chowdhury, A., Kruger, V., Chellappa, R.: Identification of humans using gait. IEEE Trans. Image Process. **13**(9), 1163–1173 (2004)
3. Sundaresan, A., RoyChowdhury, A., Chellappa, R.: A hidden markov model based framework for recognition of humans from gait sequences. In: Proceedings of 2003 International Conference on Image Processing, Barcelona, Spain, vol. 2, pp. 93–96 (2004)
4. Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., Bowyer, K.: The human ID gait challenge problem: data sets, performance, and analysis. Pattern Anal. Mach. Intell. IEEE Trans. **27**(2), 162–177 (2005)
5. Man, J., Bhanu, B.: Individual recognition using gait energy image. IEEE Trans. Pattern Anal. Mach. Intell. **28**(2), 316–322 (2006)
6. Dryden, I., Mardia, K.: Statistical shape analysis. Wiley, Chichester (1998)
7. Veeraraghavan, A., Roy-Chowdhury, A., Chellappa, R.: Matching shape sequences in video with applications in human movement analysis. Pattern Anal. Mach. Intell. IEEE Transactions on **27**(12), 1896–1909 (2005)
8. Veeraraghavan, A., Chowdhury, A., Chellappa, R.: Role of shape and kinematics in human movement analysis. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (CVPR 2004), Washington DC, USA, 27 June-2 July 2004, vol. 1, pp. I–730 – I–737 (2004)
9. Veeraraghavan, A., Chellappa, R., Roy-Chowdhury, A.: The function space of an activity. In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, vol. 1, pp. 959–968 (2006)
10. Wang, L., Ning, H., Hu, W., Tan, T.: Gait recognition based on procrustes shape analysis. In: Proceedings of the Nineth IEEE International Conference on Image Processing (ICIP, Poster), Rochester, New York, USA, vol. III (2002)
11. Cunado, D., Nash, J., Nixon, M., Carter, J.: Gait extraction and description by evidence-gathering. In: Proceedings of the International Conference on Audio and Video Based Biometric Person Authentication, vol. 48 (1995)
12. Bissacco, A., Chiuso, A., Ma, Y., Soatto, S.: Recognition of human gaits. In: Conference on Computer Vision and Pattern Recognition, Hawaii, USA, vol. 2, pp. 52–57 (2001)
13. Forner-Cordero, A., Koopman, H., van der Helm, F.: Describing gait as a sequence of states. J. Biomech. **39**(5), 948–957 (2006)
14. Liu, Z., Sarkar, S.: Improved gait recognition by gait dynamics normalization. IEEE Trans. Pattern Anal. Mach. Intell. **28**(6), 863–876 (2006)
15. Wagg, D., Nixon, M.: On automated model-based extraction and analysis of gait. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004, Seoul, Korea, pp. 11–16 (2004)
16. Kale, A., Roychowdhury, A., Chellappa, R.: Fusion of gait and face for human identification. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004,(ICASSP'04), Montreal, Quebec, Canada, vol. 5 (2004)

Gait Models for Biometrics

▶ Gait Recognition, Model-Based

Gait Recognition

- ▶ Evaluation of Gait Recognition
- ▶ Gait Biometrics, Overview

Gait Recognition, Model-Based

Chew-Yean Yam, Mark S. Nixon
University of Southampton, Southampton, UK

Synonyms

Gait models for biometrics; Knowledge-based gait recognition

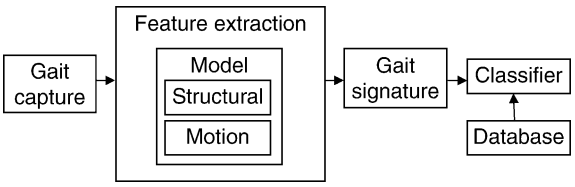
Definition

Model-based gait recognition relates to the identification using an underlying mathematical construct(s) representing the discriminatory gait characteristics (be they static or dynamic), with a set of parameters and a set of logical and quantitative relationships between them. These models are often simplified based on justifiable assumptions, e.g., a system may assume a pathologically normal gait. Such a system normally consists of gait capture, a model(s), a feature extraction scheme, a gait signature, and a classifier (Fig. 1). The model can be a 2- or 3-dimensional ▶ structural (or ▶ shape) ▶ model and/or ▶ motion model that lays the foundation for the extraction and tracking of a moving person. An alternative to a model-based approach is to analyze the motion of the human silhouette deriving recognition from the body's shape and motion. A gait signature that is unique to each person in the database is then derived from the extracted gait characteristics. In the classification stage, many pattern classification techniques can be used, such as the *k*-nearest neighbor approach.

The main advantages of the model-based approach are that it can reliably handle occlusion (especially self-occlusion), noise, scale and rotation well, as opposed to silhouette-based approaches.

Practical issues that challenge the model-based approach can be divided into two categories, which relate to the *system* and to the *person*. One of the systems-related challenges is viewpoint invariance, whilst person-related challenges include the effects of *physiological* changes (such as aging, the consistency of gait taken/enrolled at different times, whether our walking pattern changes over a longer period of time), *psychological* changes (mood), and *external factors* (load, footwear, and the physical environment).

The first model-based approach to gait biometrics was by Cunado et al. in 1997 [1, 2], featuring the ability to reliably accommodate self-occlusion and occlusion by other objects, noise, and low resolution. Also, most of the time, the parameters used within the model and their relationship to the gait are obvious, i.e., the mathematical construct may itself contain implicit/explicit meaning of the gait pattern characteristics. Though, it often suffers from high computational cost, this can be mitigated by optimization tools or increased computing power. Gait sequences are usually acquired when the subject is walking in a plane normal to the image capture device since the side view of a moving person reveals most information, though it is possible to use other views.



Gait Recognition, Model-Based. **Figure 1** Components of a typical model-based gait recognition system.

Models

In a typical model-based approach, often, a ▶ structural model and a motion model are required to serve as the basis for tracking and feature (moving human) extraction. These models can be 2- or 3- dimensional, though most of the current approaches are 2-dimensional and have shown the capability to achieve promising recognition results on large databases (>100 subjects). A structural model describes the topology or the shape of human body parts such as head, torso, hip, thigh, knee, and ankle by measurements such as the length, width, and position. This model can be made up of primitive shapes (cylinders, cones, and blobs), stick figures, or arbitrary shapes describing the edge of these body parts. On the other hand, a motion model describes the kinematics or the dynamics of the motion of each body part. Kinematics generally describe how the subject changes position with time without considering the effect of masses and forces, whereas dynamics account for the forces that act upon these body masses and the resulting motion. When developing a motion model, the constraints of gait such as the dependency of neighboring joints and the limit of motion in terms of range and direction has to be understood.

Bobick et al. used a structural model to recover static body and stride parameters (Fig. 2a) determined by the body geometry and the gait of a person [3]. Lee et al. fit ellipses to seven regions representing the human body (Fig. 2b), then derived two types of features across time: mean and standard deviation, and magnitude and phase of these moment-based region features [4].

Cunado et al. proposed an early motion-model-based approach, based on the angular motion of the hip and thigh [1, 2], where the angular motion of the hip and the thigh is described by a Fourier series. For this method, a simple structural model was used and the angular rotation as defined in Fig. 3. Although the motion model is for one leg, assuming that gait is symmetrical, the other leg can be modeled similarly, with a phase lock of ½-period shift (Fig. 4).

Cunado et al. modeled the angular motion of the thigh by

$$\theta_T = a_0 + 2\sum_{1}^{N}[b_k \cos k\omega_0 t - c_k \sin k\omega_0 t],$$

where N is the number of harmonics, ω_0 is the fundamental frequency, and a_0 is the offset. In application, the frequency data was accumulated from a series of edge-detected versions of the image sequence of the walking subject. The gait signature was derived by the multiplication of the phase and magnitude component of the Fourier description.

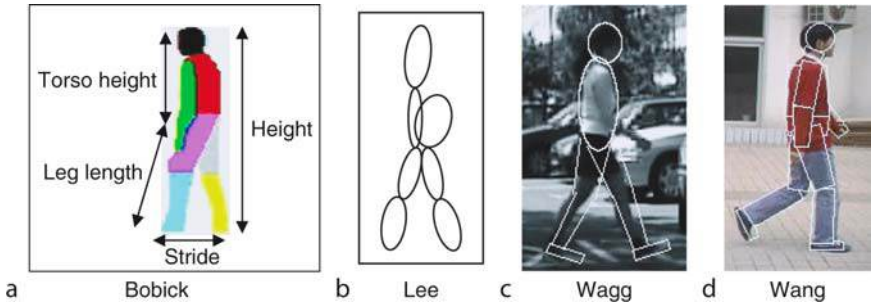
Later, Yam et al. [5] extended the approach to describe the hip, thigh, and knee angular motion of both walking and running gaits first by an empirical motion model, then by an analytical model motivated by coupled pendulum motion. Similarly, the gait signature is the phase-weighted magnitude of the Fourier description of both the thigh and knee rotation.

Bouchrika et al. [6] have proposed one of the latest motion-model-based gait feature extraction using a parametric form of elliptic Fourier descriptors to describe joint displacement.

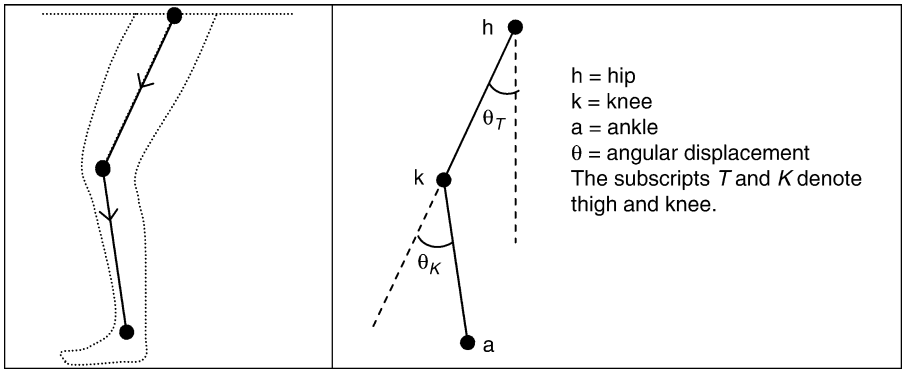
$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} X(t) * S_x \\ Y(t) * S_y \end{bmatrix},$$

where α is the angle, S_x and S_y are the scaling factors, and $X(t)$ and $Y(t)$ are Fourier summation. The joint trajectory is then fitted to the image sequence by optimizing a_0, b_0, α, S_x and S_y; the motion model fit is implemented by the Hough Transform.

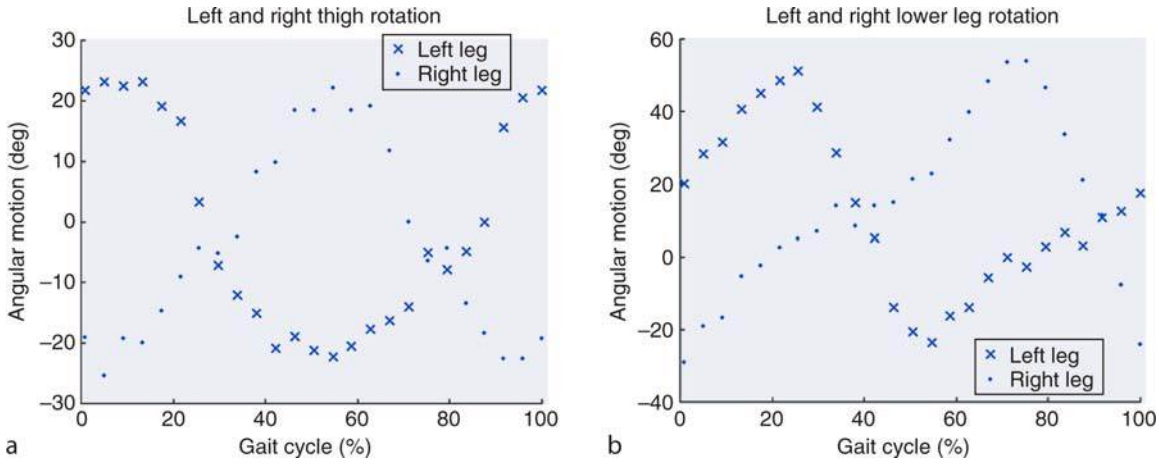
Wagg et al. (Fig. 2c) and Wang et al. (Fig. 2d) used a combination of both structural and motion models to



Gait Recognition, Model-Based. **Figure 2** Example body parameters that are used in structural models. (**a**) Bobick (**b**) Lee (**c**) Wagg (**d**) Wang.



Gait Recognition, Model-Based. **Figure 3** Structural model of a lower limb: upper and lower pendulum represents the thigh and the lower leg, respectively, connected at the knee joint.



Gait Recognition, Model-Based. Figure 4 Thigh and lower leg rotation of the left and right leg. (**a**) Left and right thigh rotation (**b**) Left and right lower leg rotation.

track and extract walking human figures [7, 8]. Wagg introduced a self-occlusion model whilst Wang used the conditional density propagation framework [9] to aid feature extraction.

Beyond the 2D models, Urtasun et al. developed a 3D gait motion model derived from a small group of subjects [10]. The joint motion is approximated by a weighted sum of the mean motion and the Eigenvectors of sample angular motion vectors. This approach also shows that it is capable of approximating running motion as well.

Feature Extraction

Feature extraction segments interesting body parts for a moving human, and extracts static and/or dynamic gait characteristics. The process normally involves model initialization, segmentation, and tracking (estimation) of the moving human from one image to the next. This is a significant step that extracts important spatial, temporal, or spatial-temporal signals from gait. Feature extraction can then be carried out in a concurrent [1, 2, 5, 8], or iterative/hierarchical [7] manner.

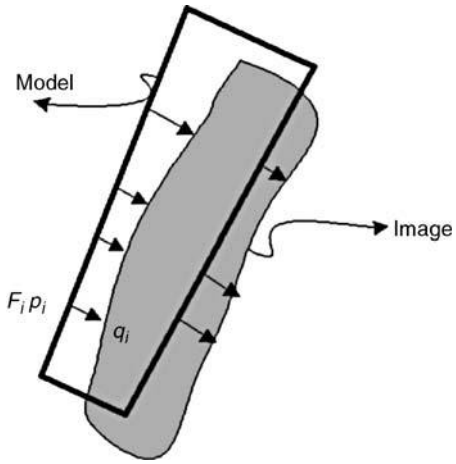
A conventional starting point of a gait cycle is the heel strike at the stance phase, although any other stage within a gait cycle can be used. Earlier techniques determine the gait cycle manually, later, many have employed automatic gait cycle detection. A gait cycle can be detected by simply identifying the stance phase; if using a bounding box method, the width of the box has the highest value during the stance phase. Other alternatives are counting the pixels of the human figure, using binary mask (Fig. 5) by approximating the outer region of the leg swing [7].

Quality of Feature Extraction

A good model configuration is defined as one that yields a high correlation between the model and the subject's image. Useful measures for computing model and image data correlation include *edge correspondence* and *region correspondence* [8]. Edge correspondence is a measure of how closely model edges coincide with image edges, whilst region correspondence is a measure of similarity between the image region enclosed by the model and that corresponding to the image of the subject. These two measures are used together. A high edge correspondence indicates that the model is closely aligned with image edges; however, it does not guarantee that the model matches the correct edges. If the initial model configuration is poor, or the subject is occluded, the match may be coincidental. For this reason, region correspondence is also required.

Another measure is a pose evaluation function (PEF) which combines the boundary (edge) matching error and the region matching error to achieve both accuracy and robustness. For each pixel, p_i, in the boundary of the projected human model, the corresponding pixel in the edge image along the gradient

Gait Recognition, Model-Based. **Figure 5** Binary mask to detect gait cycle. The sum edge strength within the mask varies periodically during the subject's gait and the heel strike being the greatest.



Gait Recognition, Model-Based. **Figure 6** Measuring the boundary matching error.

direction at p_i (Fig. 6) is searched. In other words, the pixel nearest to p_i and along that direction is desired. Given that q_i is the corresponding pixel and that $\boldsymbol{F}_i$ stands for the vector $\overline{p_i q_i}$, the matching error of pixel p_i to q_i can be measured as the norm $\|\mathbf{F}_i\|$. Then the average of the matching errors of all pixels in the boundary of the projected human model is defined as the boundary matching error

$$E_b = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{F}_i\|,$$

where N is the number of the pixels in the boundary.

In general, the boundary matching error measures the similarity between the human model and image data, but it is insufficient under certain circumstances, as illustrated in Fig. 7a, where a model part falls into the gap between two body parts in the edge image. Although it is obviously badly-fitted, the model part may have a small boundary matching error. To avoid such ambiguities, region information is further considered. Figure 7b illustrates the region matching. Here the region of the projected human model that is fitted into the image data is divided into two parts: P_1 is the model region overlapped with the image data and P_2 is the rest of the model region. Then the matching error with respect to the region information is defined by

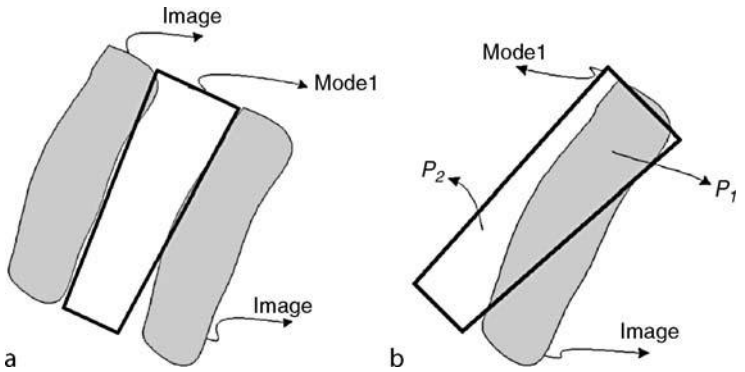
$$E_r = |P_2|/(|P_1| + |P_2|)$$

where $|P_i|$, $(i = 1, 2)$ is the area, i.e., the number of pixels in the corresponding region.

Recognition

A gait signature is a discriminatory feature vector that can distinguish individual. These signatures have invariant properties embedded in a person such as stride length, person's height/width, gait cycle and self-occlusion, and that related to the imaging system such as translation, rotation, scale, noise, and occlusion by other objects. These signatures can be of static [3], dynamic [2, 5] or a fusion of static and dynamic [7, 8] characteristics of gait or with other biometrics [11, 12]. The fusion can happen either at the feature extraction stage or at the classification stage. On the Southampton datasets of 115 subjects filmed indoors (in controlled conditions) and outdoors (with effects of shadows, background objects, and changing illumination) Wagg's approach achieved an overall CCR of 98.6% on the indoor data and 87.1% on the outdoor data.

In the case of 3D approach [10], experiments show that the first six coefficients of that motion model can



Gait Recognition, Model-Based. **Figure 7** Illustrating the necessity of simultaneous boundary and region matching. (**a**) A typical ambiguity: a model part falls into the gap between two body parts (**b**) Measuring region matching error.

characterize 90% gait patterns of the database used. This resulted in a very compact gait signature, which requires only the first three coefficients to form separate clusters for each subject. It is interesting that this study found that the first few coefficients could represent physiological characteristics like weight, height, gender or age, while the remaining ones can be used to distinguish individual characteristics. Another interesting finding is that the nature of the gait signature for running derived from this 3D motion model is similar to that of Yam et al., that is, signature clusters are more dispersed within subject, and span more widely within the signature space, as compared to that of walking. Both studies were based on data collected by having subjects running on the treadmill.

Conclusions and Outlook

Using a model is an appealing way to handle known difficulty in subject acquisition and description for gait biometrics. There is a selection of models and approaches which can handle walking and running. Clearly, the use of a model introduces specificity into the feature extraction and description process, though this is generally at the cost of increased computation. Given their advantages, it is then likely that model-based approaches will continue to play a part in the evolution of systems which deploy gait as a biometric. Currently, practical advantages of three-dimensional (3D) approaches have yet to be explored and investigated. Given that human motion occurs in space and time, it is likely that much information is embedded within the 3D space. Further, 3D approaches may provide a more effective way to handle issues like occlusion, pose, and view point. Therefore, 3D model-based gait recognition may be a good way to move forward.

Related Entries

- ▶ Gait Recognition, Model-Based
- ▶ Human Detection and Tracking
- ▶ Markerless 3D Human Motion Capture from Images
- ▶ Multibiometrics
- ▶ Silhouette-Based Recognition

References

1. Cunado, D., Nixon, M.S., Carter, J.N.: Using gait as a biometric, via phase-weighted magnitude spectra. In: First International Conference on Audio- and Video-Based Biometric Person Authentication, pp. 95–102 Crans-Montana, Switzerland, (1997)
2. Cunado, D., Nixon, M.S., Carter, J.N.: Automatic extraction and description of human gait models for recognition purposes. Comput. Vis. Image Underst. **90**(1), 1–41 (2003)
3. Bobick, A.F., Johnson, A.Y.: Gait recognition using static, activity-specific parameters. In: Proceedings of IEEE Computer Vision and Pattern Recognition Conference (CVPR'01), pp. 423–430 Kauai, Hawaii, (2001)
4. Lee, L., Grimson, W.E.L.: Gait analysis for recognition and classification. In: Proceedings of Automatic Face and Gesture Recognition, pp. 148–155 Washington, DC, (2002)
5. Yam, C.Y., Nixon, M.S., Carter, J.N.: Automated person recognition by walking and running via model-based approaches. Pattern Recognit. **37**, 1057–1072 (2004)
6. Bouchrika, I., Nixon, M.S.: Model-based feature extraction for gait analysis and recognition. In: Mirage: Computer Vision/Computer Graphics Collaboration Techniques and Applications, INRIA Rocquencourt, France, March (2007)

7. Wagg, D.K., Nixon, M.S.: On automated model-based extraction and analysis of gait. In: Proceedings of Sixth International Conference on Automatic Face and Gesture Recognition, pp. 11–16 Seoul, South Korea, (2004)
8. Wang, L., Tan, T., Ning, H., Hu, W.: Fusion of static and dynamic body biometrics for gait recognition. IEEE Trans. Circuits Syst. Video Technol. (Special Issue on Image- and Video-Based Biometrics) **14**(2), 149–158 (2004)
9. Isard, M., Blake, A.: CONDENSATION – conditional density propagation for visual tracking. Int. J. Comput. Vis. **29**(1), 5–28 (1998)
10. Urtasun, R., Fua, P.: 3D tracking for gait characterization and recognition. In: Proceedings of Sixth IEEE International Conference on Automatic Face and Gesture recognition, Seoul, South Korea, pp. 17–22 (2004)
11. Kale, A., RoyChowdhury, A.K., Chellappa, R.: Fusion of gait and face for human identification. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Canada, May 2004, vol. 5, pp. 901–904 (2004)
12. Shakhnarovich, G., Lee, L., Darrell, T.: Integrated face and gait recognition from multiple views. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 439–446 Hawaii, (2001)

Gait Recognition, Motion Analysis for

Ahmed Elgammal
Department of Computer Science, Rutgers University, Piscataway, NJ, USA

Synonyms

Appearance-based gait analysis; Silhouette analysis for gait recognition

Definition

The appearance of gait in an image sequence is a spatiotemporal process that characterizes the walker. The spatiotemporal characteristics of gait contain rich perceptual information about the body configuration, the person's gender, the person's identity, and even the emotional states of the person. Motion analysis for gait recognition is a computer vision task that aims to capture discriminative spatiotemporal features (signature) from image sequences in order to achieve human identification. Such a signature ought to be invariant to the presence of various viewing conditions, such as viewpoint, people clothing, etc. In contrast to Model-based gait analysis systems, which is another article, the goal here is to capture gait characteristics without fitting a body model or locating the body limbs, rather by analyzing the feature distribution over the space and time extent of the motion.

Human Gait as a Biometric

Human gait is a valuable biometric cue that has the potential to be used for human identification similar to other biometric features, such as faces and fingerprints. Gait has significant advantages compared to other biometric features since it is easily observable in an unintrusive way, it does not require collaborative subjects, and it is difficult to disguise [1]. Therefore, using gait as a biometric feature has a great potential for human identification in public places for surveillance and for security. A fundamental challenge in gait recognition is to develop robust algorithms that can extract visual gait features invariant to the presence of various conditions that affect people's appearance, as well as conditions that affect people's gait. That includes, viewpoint, clothing, walking surface, shoe type, object carried, etc. [2].

Johansson's seminal psychophysical experiments [3] showed that humans can recognize biological motion, such as gait, from Moving Light Displays (MLD). Cutting and Kozlowski [4] showed that humans can also identify friends from their gait using MLD. Motivated by these results, many researchers in different disciplines, have shown that the spatiotemporal characteristics of gait contain rich perceptual information about the body configuration, the person's gender, the person's identity, and even the emotional states of the person. That motivated extensive recent computer vision research on extracting features from gait.

Vision-based human motion tracking and analysis systems have promising potentials for many applications, such as visual surveillance in public area, activity recognition, sport analysis, video retrieval, and human–computer interaction. Extensive research has been done in this area in the last two decades with lots of promising results. For excellent literature surveys in the subject, the reader can refer to [5, 6]. The human body is an articulated object with a large number of degrees of freedom. This fact makes the problems of tracking the

body configuration and extracting biometrics very challenging. Besides the articulation nature of the body, the variability in people's appearance adds to the problems. Human gait is a special case of the general problem of human motion analysis, and to some extent, is easier. This is because of the physical constraints on such a motion as well as the periodic nature of it.

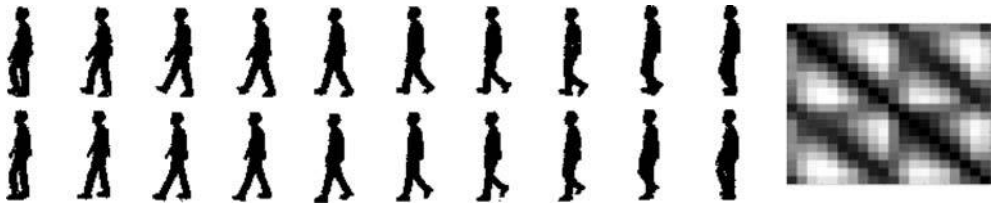
The appearance of gait in an image sequence is a spatiotemporal process that characterizes the walker. Gait recognition algorithms, generally, aim to capture discriminative spatiotemporal features (signature) from image sequences in order to achieve human identification. Gait analysis approaches can be categorized according to the way the gait features are extracted for classification. There are two broad categories of approaches: model-based approaches and appearance-based approaches. Model-based approaches, e.g., [1], fit 3D body models or intermediate body representations to body limbs in order to extract proper features (parameters) that describe the dynamics of the gait (see the related entry on "Model-based Gait Recognition" for details). Model-based approaches typically require a large number of pixels on the tracked target to fit their model, i.e., high resolution zoomed-in images are required on the tracked person. In contrast, appearance-based approaches aim to capture a spatiotemporal gait characteristic directly from input sequences without fitting a body model. The appearance-based approaches are mainly motivated by the psychophysical experiments, mentioned earlier, e.g., [3, 4], which showed that spatiotemporal patterns such as Moving Light Displays could capture important gait information without the need of finding limbs. Appearance-based approaches do not require high resolution on subjects, which makes them more applicable in outdoor surveillance applications where the subjects can be at a large distance from the camera.

Characteristics and Challenges of Gait Motion

Gait is a 3D articulated periodic motion that is projected into 2D image sequences. Therefore, the appearance of a gait motion in an image sequence is a spatiotemporal pattern, i.e., a spatial distribution of features that changes over time. Researchers have developed several algorithms for capturing gait signature from such spatiotemporal patterns by looking at the space-time volume of features. The observed shapes of the human body, in terms of the occluding contours of the body (silhouettes), are examples of such spatiotemporal patterns, which contain rich perceptual information about the body configuration, the motion performed, the person's gender, the person's identity, and even the emotional states of the person. Objects occluding contours, in general, have a great role in perception [7] and have been traditionally used in computational vision, besides other appearance cues, to determine object category and pose.

The objective of any gait tracking and analysis system is to track the global deformations of contours over time and to capture invariant gait signature from such contours. There are several challenges to achieve this goal. An observed person's contour in a given image is a function of many factors, such as the person's body build (tall, short, big, small, etc.), the body configuration, the person's clothing and the viewpoint. Such factors can be relevant or irrelevant depending on the application. Modeling these sources of variabilities is essential to achieve successful trackers and to extract gait biometric features. Modeling the human body dynamic shape space is hard, since both the dynamics of shape (different postures) and the static variability in different people's shapes have to be considered. Such shape space lies on a nonlinear ▶ manifold.

Figure 1 shows an example of a walking cycle from a side view where each row shows half a walking cycle. The shapes during a gait cycle temporally undergo deformations and self-occlusion. The viewpoint from which the gait is captured imposes self-similarity on the observed shapes over time. This similarity can be noticed by comparing the corresponding shapes at the two rows in Fig. 1. This right part of the figure shows the correlation between these shapes. The similarity between the corresponding shapes in the two half cycles is exhibited by the dark diagonally parallel bands in the correlation plot. The similarity in the observed shapes indicates a nonlinear relation between the observed gait and the kinematics of the gait. This can be noticed by closely inspecting the two shapes in the middle of the two rows in Fig. 1. These two shapes correspond to the farthest points in the walking cycle kinematically (the top has the right leg in front while the bottom has the left leg in front). In the Euclidean visual input space (observed shapes) these two points are very close to each other as can be noticed from the distance plot on the right of Fig. 1. This nonlinear relation between the observed shapes



Gait Recognition, Motion Analysis for. **Figure 1** Twenty sample frames from a walking cycle from a side view. Each row represents half a cycle. Notice the similarity between the two half cycles. The right part shows the similarity plot: each row and column of the plot corresponds to one sample. Darker means closer distance and brighter means larger distances. The two dark lines parallel to the diagonal show the similarity between the two half cycles.

and the kinematics poses a problem to gait tracking and analysis systems. However, such similarity can be useful in extracting gait features. For example, the temporal self-similarity characteristic has been exploited in the work of BenAbdelkader et al. [8] for gait recognition.

Extracting Gait Signature from Motion

There have been extensive research on appearance-based extraction of gait signatures. Typical preprocessing steps for gait analysis include detecting and tracking the human subject in order to locate a bounding box containing the motion and/or extracting the body silhouette (see the related entry on human detection and tracking).

One of the early papers on gait analysis using spatiotemporal features is the work of Niyogi and Adelson [9] where a spatiotemporal pattern (corresponding to leg motion) was used to detect gait motion in an image sequence represented as an XYT volume. Gait was then parameterized with four angles for recognition. Murase and Sakai [10] used a parametric eigenspace representation to represent a moving object using Principle Component Analysis (PCA). In their work, the extracted silhouettes were projected into an eigenspace where a walking cycle forms a closed trajectory in that space. Spatiotemporal correlations between a given trajectory and a database of trajectories were used to perform the recognition. Huang et al. [11] extended the method using Canonical space transformation (CST) based on Canonical Anaylsis (CA), with eigenspace transformation for feature extraction.

Little and Boyd [12] exploited the spatial distribution of optical flow to extract spatiotemporal features. From dense optical flow, they extracted scale-independent features capturing the spatial distribution of the flow using moments. This facilitates capturing the spatial layout of the motion, or as they call it "the shape of the motion." Periodicity analysis was then done on these features to capture gait signatures for recognition. BenAbdelkader et al. [8] used image self-similarity plots (similar to Fig. 1) to capture the spatiotemporal characteristics of gait. Given bounding boxes around a tracked subject, correlation is used to measure self-similarity between different time frames in the form of similarity plots. PCA analysis was used to reduce the dimensionality of such similarity plots for recognition. Hayfron-Acquah et al. [13] used spatial symmetry information to capture gait characteristics from silhouettes. Given a walking cycle, a symmetry operator was used to extract a symmetry map for each silhouette instance in the cycle. Fourier transform was used to extract descriptors from such symmetry maps for recognition.

Since gait is a temporal sequence, researchers have investigated the use of Hidden Markov Models (HMM) to represent and capture gait motion characteristics. HMMs have been successfully used in many speech recognition systems, as well as gesture recognition applications. Typically a left-right HMM with a small number of states (three to five) is sufficient to model the gait of each subject in the database, where the HMMs are trained from features extracted from silhouettes. In [14], HMM was used to capture gait dynamics from quantized Hu moments of silhouettes. HMM was also used in [15] with features representing silhouette width distribution.

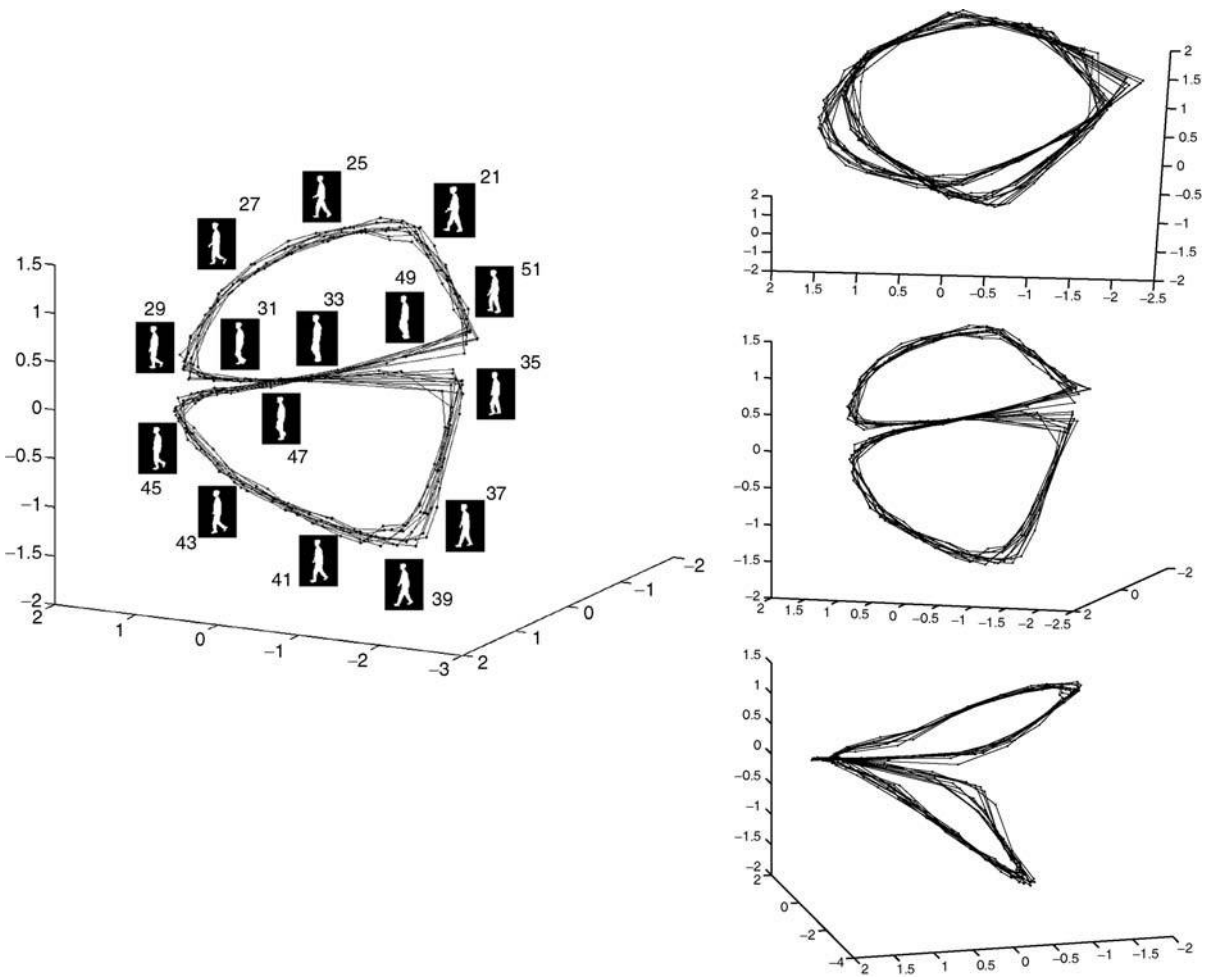
Lee and Elgammal [16] used bilinear and multilinear models to factorize the spatiotemporal gait process into gait style and gait content factors. A nonlinear mapping was learned from a unit circle (representing a gait cycle) to the silhouettes' shape space. The unit circle represents a unified model for the gait manifold

of different people, therefore, any spatiotemporal characteristics of the gait of a specific person should exist on the mapping space. Bilinear and multilinear models were used to factorize such mapping to extract gait signatures.

Manifold-based Representation for Gait Analysis

Despite the high dimensionality of the human body configuration space, any body motion is constrained by the physical dynamics, body constraints, and the motion type. Therefore, many human activities lie intrinsically on low dimensional manifolds. This is true for the body kinematics, as well as for the observed motion through image sequences. For certain classes of motion like gait, facial expression, and simple gestures, considering a single person and factoring out other sources of variability, the deformations will lie on a one-dimensional manifold. Recently many researchers have developed techniques and representations for gait analysis that exploit such manifold structure, whether in the visual space or in the kinematic space, e.g. [17, 18]. Modeling the gait manifold was earlier used for gait recognition in [10].

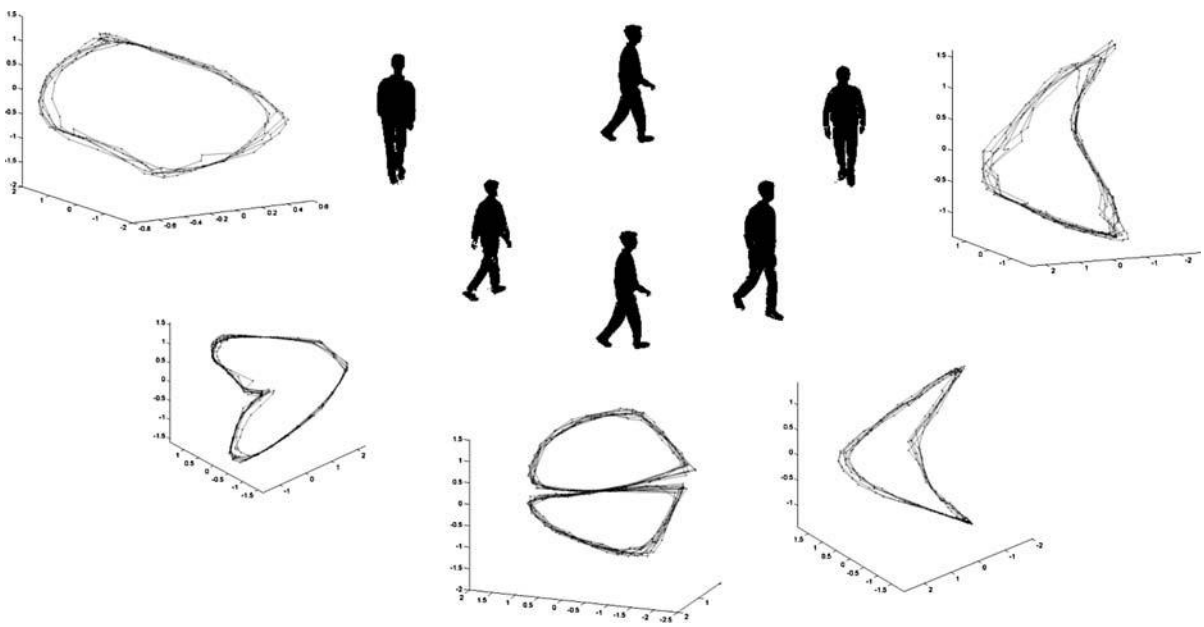
Intuitively, the gait is a one-dimensional closed manifold that is embedded in a high dimensional visual space. Such a manifold can twist and self-intersect in such high dimensional visual space. This can be



Gait Recognition, Motion Analysis for. **Figure 2** Embedded gait manifold for a side view of the walker. *Left*: sample frames from a walking cycle along the manifold with the frame numbers shown to indicate the order. Ten walking cycles are shown. *Right*: three different views of the manifold. © IEEE.

noticed by considering the human silhouette through the walking cycle, (as shown in Fig. 1) as points in a high dimensional visual input space. Given the spatial and the temporal constraints, it is expected that these points will lay on a closed trajectory. In order to achieve a low dimensional embedding of the gait manifold (► manifold embedding), dimensionality reduction techniques can be used. Linear dimensionality reduction can be used to achieve an embedding, as in [10]. However, in such a case the two half cycles would be collapsed to each other because of the similarity in the shape space. Nonlinear dimensionality reduction techniques such as LLE [19], Isomap [20], GPLVM [21], and others can successfully embed the gait ► manifold in a way that separates the two half cycles. As a result of nonlinear dimensionality reduction, an embedding (and a visualization) of the gait manifold can be obtained in a low-dimensional Euclidean space [17]. Figure 2 illustrates an example embedded manifold for a side view of the walker. The data used are from the CMU Mobo gait data set which contains 25 people from six different view points. Data sets of walking people from multiple views are used in this experiment. Each data set consists of 300 frames and each containing about 8–11 walking cycles of the same person from a certain view points. The walkers were using treadmill which might result in different dynamics from the natural walking. Figure 3 illustrates the embedded manifolds for five different view points of the walker. For a given view point, the walking cycle evolves along a closed curve in the embedded space, i.e., only one degree of freedom controls the walking cycle, which corresponds to the constrained body pose as a function of the time. Such a conclusion conforms to the intuition that the gait manifold is one-dimensional.

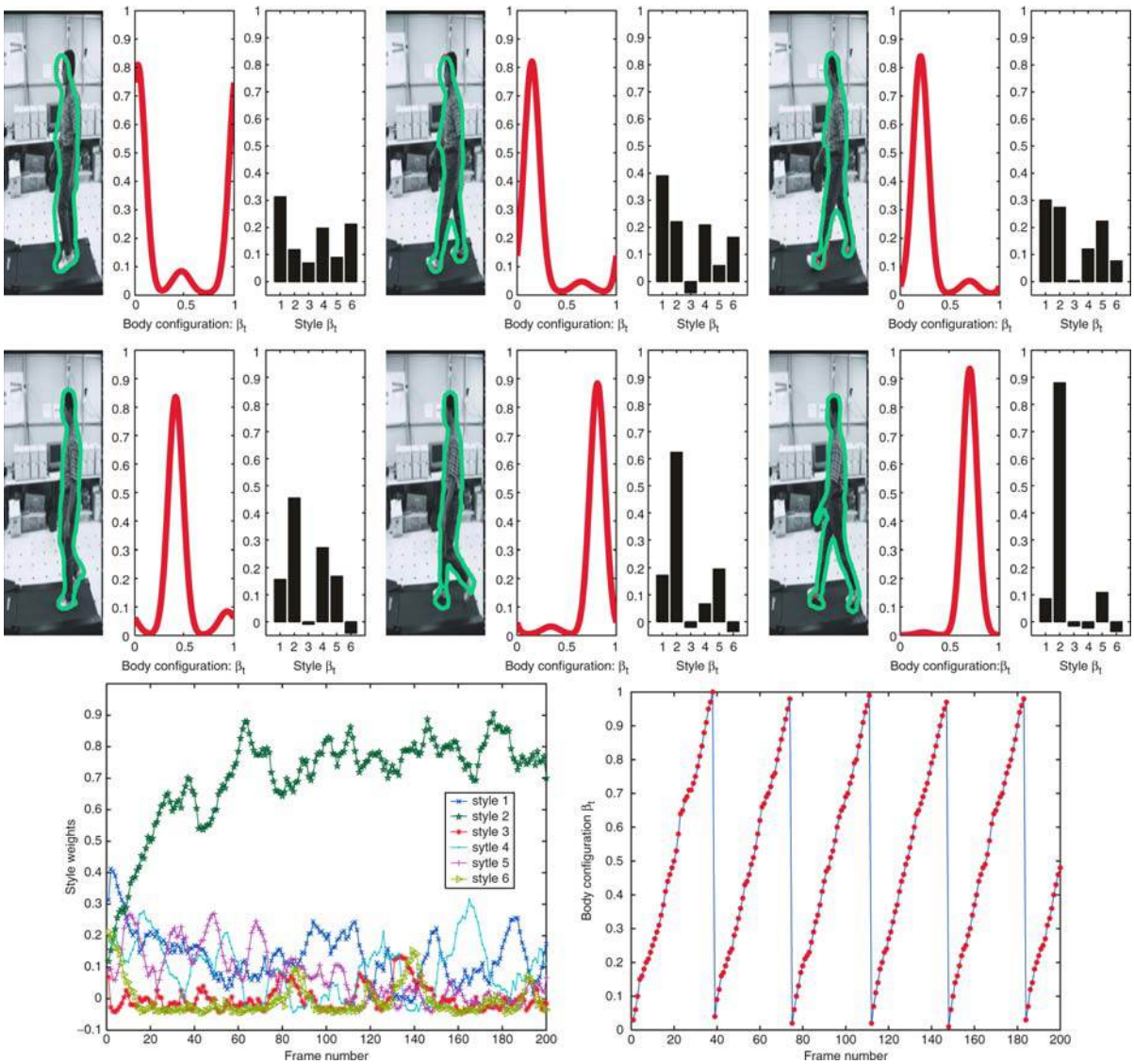
As can be noticed in Fig. 3, The manifold twists in the embedding space given the different viewpoints, which impose different self occlusions. The least twisted manifold is the manifold for the back view as this is the least self occluding view (left most manifold in Fig. 3. In this case the manifold can be embedded in a two dimensional space. For other views, the curve starts to twist to be a three-dimensional space curve. This is primarily because of the similarity imposed by the view point which attracts far away points on the manifold closer. The ultimate twist happens in the side view manifold where the curve twists to get the shape of the numeral 8 where each cycle of the eight (half eight) lies in a different plane. Each half of the "eight" figure corresponds to half a walking cycle. The cross point represents the body pose where it is



Gait Recognition, Motion Analysis for. Figure 3 Embedded manifolds for five different views of the walkers. Frontal view manifold is the right most one and back view manifold is the leftmost one. The view of the manifold that best illustrates its shape in the 3D embedding space is visualized. © IEEE.

totally ambiguous from the side view to determine from the shape of the contour which leg is in front, as can be noticed in Fig. 2. Therefore, in a side view, a three-dimensional embedding space is the least that can be used to discriminate the different body poses. Embedding a side view cycle in a two-dimensional embedding space results in an embedding similar to that shown in top right of Fig. 2 where the two half cycles lie over each other. Interestingly, despite that the side view is the most problematic view of the gait, most gait recognition systems seem to favor such view for recognition! Different people are expected to have different manifolds. However, such manifolds are all topologically equivalent.

The example embeddings shown here are for silhouette data, i.e., the *visual manifold* of the gait is



Gait Recognition, Motion Analysis for. Figure 4 Adaptive Contour Tracking of Gait: (**a**) tracking through sample frames. (**b**) adapting to the target style. (**c**) the tracked body configuration showing a constant speed dynamic system. From [22].

embedded. Similar embedding can be obtained for kinematic data, in such a case the *kinematic manifold* of the gait is embedded. In such a case PCA would be sufficient to achieve an embedding. The importance of such embedded representations is that they provide a low dimensional representation for tracking the gait motion. Only a one-dimensional parameter is needed to control and track the gait motion. This leads to a simple constant speed dynamic model for the gait. Figure 4 shows an example of gait contour tracking system [22] that uses an embedded representation of the gait manifold. As a result, a constant speed linear dynamics is achieved (Fig. 4b). The tracker can also adapt to the tracked person shape style and identify that style from a database of styles (Fig. 4c).

Explicit manifold representation for gait is not only useful for tracking and pose estimation, but also can be used in gait recognition systems. Different people are expected to have different manifolds for the appearance of their gait. However, such manifolds are all topologically equivalent to a unit circle. A person's gait manifold can be thought of as a twisted circle in the input space. The spatiotemporal process of gait is captured in the twist of a given person's manifold. Therefore, a person's gait signature can be captured by modeling how a unit circle (an ideal manifold) can deform to fit that person's gait manifold. This can be achieved by fitting a nonlinear warping function between a unit circle and a given person's silhouette sequence. In [23] this approach was used to capture gait signatures by factorizing the warping functions' coefficient space to obtain a low-dimensional gait signature space for recognition.

Summary

Appearance-based analysis of gait is motivated and justified by psychophysical experiments. Appearance-based approaches for gait recognition aim to extract a gait signature from the spatial and temporal distribution of the features on a tracked subject without the need to fit a body model or to locate limbs. Such approaches have proved very successful in gait recognition and are applicable in scenarios where the gait biometric features can only be extracted from a distance. There are many limitations to the current gait recognition systems including achieving invariant to viewing conditions, such as viewpoint invariant. Recent progress in manifold-based representation of gait, as well as factorized models, such as multilinear tensor models provides potential solutions to such problems.

Related Entries

- ▶ Gait Recognition, Model-Based
- ▶ Human detection and tracking

References

1. Cunado D., Nixon M.S., Carter J.: Automatic extraction and description of human gait models for recognition purposes. Compt. Vision Image Understand. **90**, 1–41 (2003)
2. Sarkar S., Phillips P.J., Liu Z., Vega I.R., Grother P., Bowyer K.W.: The humanid gait challenge problem: Data sets, performance, and analysis. IEEE Trans. PAMI **27**(2), 162–177 (2005)
3. Johansson G.: Visual motion perception. Sci. Am. **232**, 76–88 (1975)
4. Cutting J.E., Kozlowski L.T.: Recognizing friends by their walk: gait perception without familiarity cues. Bull. Psychonomic Soc. **9**(5), 353–356 (1977)
5. Gavrila D.M.: The visual analysis of human movement: a survey. Comput. Vis. Image Understand. **73**(1), 82–98 (1999). DOI http://dx.doi.org/10.1006/cviu.1998.0716
6. Moeslund T.B., Hilton A., Krüger V.: A survey of advances in vision-based human motion capture and analysis. Comput. Vis. Image Underst. **104**(2), 90 –126 (2006). DOI http://dx.doi.org/10.1016/j.cviu.2006.08.002
7. Palmer, S.E.: Vision Science, Photons to Phenomenology. MIT, Cambridge, MA (1999)
8. BenAdbelkader, C., Cutler, R., Davis, L.: Motion-based recognition of people using image self-similarity. In: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 254–259 (2002)
9. Niyogi, S., Adelson, E.: Analyzing and recognition walking figures in xyt. In: Proceedings of IEEE CVPR, pp. 469–474 (1994)
10. Murase H., Sakai R.: Moving object recognition in eigenspace representation: gait analysisand lip reading. Pattern Recog. Lett. **17**, 155–162 (1996)
11. Huang P., Haris C., Nixon, M.: Recogising humans by gait via parametric canonical space. Artif. Intell. Eng. **13**, 359–366 (1999)
12. Little, J.J., Boyd, J.E.: Recognizing people by their gait: The shape of motion. Videre: J. Comput. Vision Res. **1**(2) (1998)
13. Hayfron-Aquah J.B., Nixon M.S., Carter J.N.: Automatic gait recognition by symmetry analysis. Pattern Recogn. Lett. **24**, 2175–2183 (2003)
14. He, Q., Debrunner, C.: Individual recognition from periodic activity using hidden markov models. In Proceedings of the Workshop on Human Motion (Humo) (Dec, 2007). HUMO. IEEE Computer Society, Washington, DC
15. Kale, A., Sundaresan, A., Rajagopalan, A.N., Cuntoor, N.P., Roy-Chowdhury, A.K., Kruger, V., Chellappa, R.: Identification

of human using gait. IEEE Trans. Image Process. **13**(9), 1163–1173 (2004)
16. Lee, C.S., Elgammal, A.: Gait style and gait content: Bilinear model for gait recogntion using gait re-sampling. In: Proceedings of FGR, pp. 147–152 (2004)
17. Elgammal, A., Lee, C.S.: Inferring 3d body pose from silhouettes using activity manifold learning. In: Proceeedings of CVPR, vol. 2, pp. 681–688 (2004)
18. Urtasun, R., Fleet, D.J., Hertzmann, A., Fua, P.: Priors for people tracking from small training sets. In: ICCV, Beijing, China pp. 403–410 (2005)
19. Roweis S., Saul L.: Nonlinear dimensionality reduction by locally linar embedding. Science **290**(5500), 2323–2326 (2000)
20. Tenenbaum, J.: Mapping a manifold of perceptual observations. In: Proceedings of Advances in Neural Information Processing (NIPS), vol. 10, pp. 682–688 (1998)
21. Lawrence, N.D.: Gaussian process models for visualisation of high dimensional data. In: Proceedings of NIPS. BMVC, Oxford, UK (2004)
22. Lee, C.S., Elgammal, A.: Style adaptive bayesian tracking using explicit manifold learning. In: Proceedings of British Machine Vision Conference (2005)
23. Lee, C.S., Elgammal, A.: Audio- and Video-based Biometric Person Authentication Conference AVBPA. Terrytown, NY, USA (2005)

Gait Recognition, Silhouette-Based

Jeffrey E. Boyd[1], James J. Little[2]
[1]University of Calgary, Calgary, AB, Canada
[2]University of British Columbia, Vancouver, BC, Canada

Definition

Silhouette-based gait recognition is the analysis of walking human figures for the purpose of biometric recognition. Gait biometrics offers the advantage of covertness; acquisition is possible without the awareness or cooperation of the subject. The analysis may apply to a single static image, or to a temporal sequence of images, i.e., video.

Introduction

The phenomenon of gait is the "coordinated, cyclic combination of movements that result in human locomotion" [1]. Gait is necessary for human mobility and is therefore ubiquitous and easy to observe.

The common experience of recognizing a friend from a distance by the way they walk has inspired the use of gait as a biometric feature. In fact, Cutting and Kozlowski [2], using ▶ moving light displays to isolate the motion stimulus, demonstrated that humans can indeed identify familiar people from gait. In their experiments, seven subjects identified the gaits of a subset of six subjects correctly at a rate of 38%. While this rate is less than adequate for biometrics, it is significantly better than random (17% in for their sample size), and validates the human source of inspiration.

To convert a gait into a feature vector suitable for biometrics, one can *characterize the motion* in the gait, e.g., by analyzing joint angles and limb trajectories, or by measuring the overall pattern of motion. Alternatively, one can *measure critical body dimensions* such as height or limb lengths. In the later approach, biometric features can be measured statically, but the motion in the gait provides a convenient mechanism to reveal joint positions, and consequently, limb lengths. McGeer's work on *passive dynamic walkers* [3, 4] reveals the extent to which gait motion relates to body mass and limb lengths: in the passive dynamic model of a human gait, the motion is a stable limit cycle that is a direct result of body mass and limb length. Factors not accounted for in McGeer's original model are muscle activation (gravity powers a passive dynamic walker), walking surface, injury, and fatigue. Intuitively, the motion in a gait is a reflection of the mass and skeletal dimensions of the walker. McGeer's passive dynamic model leads to more sophisticated models that account for some of these other factors. For example, see the work of Kuo [5, 6].

Confounding factors in gait biometrics include clothing and footwear. Clothing can change the observed pattern of motion and make it difficult to accurately locate joint positions. The effect of footwear is more complex. Some variation in footwear causes changes in muscle activation, but causes no outwardly visible change in the pattern of motion [7], whereas other footwear changes will alter gait.

▶ Silhouette-based gait recognition extracts the form of a walking subject, and then computes a feature vector that describes either the pattern of motion in the gait, or the physical dimensions of the subject. A classifier then matches the feature vector against previously acquired examples for identification or verification.

Silhouettes

Definitions of silhouette are often ambiguous: some definitions refer to the region covered by a figure, whereas other definitions refer to the boundary between a figure and its background. In the context of silhouette-based gait biometrics, we assume that the silhouette refers to the region, rather than the border. Nevertheless, there are related examples that use the boundary, e.g., see Baumberg and Hogg [8].

To form a silhouette of a walking figure requires the ▶ segmentation of image pixels into foreground (the moving figure) and background (everything else) sets of pixels. The silhouette is the set of foreground pixels. The easiest way to acquire a reliable silhouette is *chroma-keying* [9], which relies on color disparities between a backdrop and the foreground subject. The background color (usually green or blue), is chosen to make the color discrimination robust. Figure 2d shows an example of chroma-keying in gait analysis. The unusual color of the backdrop makes the subject aware that they are under surveillance, negating the covertness of gait biometrics.

▶ Background subtraction obviates the need for a colored backdrop by measuring the naturally occurring scene behind the subject. This entails estimating the statistical properties (usually in the luminance and color) of every pixel over one or more frames of video. By comparing the background estimate with subsequent frames of video, one can classify foreground pixels as those that do not match the background. The classifier can be as simple as thresholding of the absolute difference between the background and video frames. In most cases, the background estimation and subtraction are merged into an online system that continuously computes pixel differences and then updates the background for each frame of video. Background subtraction requires that the background and camera be stationary. Stauffer and Grimson [10] describe a widely used background subtraction method that uses a multimodal estimate of background statistics to produce reliable silhouettes of moving objects. Their method is robust in the presence of some background motions (e.g., rustling leaves or swaying tree branches).

The projection of motion in a scene onto a camera image plane is called a motion field. When a human figure is walking, segmenting moving from slow or stationary pixels in the motion field will extract a silhouette of the figure. Additionally, a motion field provides richer information than a simple silhouette because it indicates not only where the subject is moving, but also how fast the various body parts are moving. In general, it is not possible to measure a motion field, but one can measure ▶ optical flow, an approximation to the motion field that is sufficient for biometric gait recognition. If one imagines the luminance of pixels to be a fluid that can flow around an image, the optical flow estimates the movement of that fluid. It is, in part, related to the motion field, but is not necessarily equal to the motion field in all cases. Barron et al. [11] provide a comparative survey of some well-known optical flow algorithms. For example, see Fig. 2a.

Most silhouette-based biometric gait analysis focuses on a view of the subject orthogonal to the sagittal plane of the subject, i.e., the subject walks across the field of view rather than toward or away from the camera. We believe that this preference exists because front or rear views of the subject show mostly side-to-side motion and do not reveal either joint location or the complex patterns of limb motion.

Marker-based motion capture, e.g., Johansson's moving light displays, offers a counterpoint to silhouettes that are less practical for biometrics, but are useful for gaining insight into the perceptual issues surrounding gait [12, 13].

Duration of Observation

In general, it is desirable to observe the gait as long as possible. One way to extend the duration of an observation indefinitely is to have a subject walk on a treadmill in front of a stationary camera, e.g., see Fig. 2b. However, this requires the cooperation and awareness of the subject.

Alternatively, allowing the camera to pan with the motion of the subject can extend the observation time without the subject walking on a special apparatus. However, when the camera moves, the images acquired contain both the movement of the subject, and the background. The changing background makes accurate background subtraction difficult.

Using a static camera simplifies both the apparatus and the processing to extract the silhouette, but the duration of observation is limited by the time it takes the subject to cross the field of view of the camera. The actual duration will vary with the angular width of

the field of view, the distance between the subject and the camera, and the speed of the subject. The practical limit on distance to subject depends on the resolution of the camera. Higher resolutions allow the subject to be further away while maintaining enough pixel coverage to measure biometric feature vectors accurately. In examples reported in the literature that use a static cameras and subjects walking on the ground, the typical duration of observation is approximately three to six strides.

Periodicity and Synchronization

Gait is a periodic phenomenon, so the silhouette of a walker varies with position in the gait cycle. Consequently, it is necessary to synchronize measurements of the silhouette to positions in the gait cycle. In turn, this requires measurement of the frequency of the gait and establishment of a phase reference within the gait cycle.

The method used to perform the synchronization depends on the particular measurements acquired and can serve to differentiate gait analysis methods. For example, Little and Boyd [14] measure the frequency from the oscillations of the centroid of the figure. To establish a phase reference, they use the phase of an oscillating measurement. In methods that measure height, e.g., Ben-Abdelkader et al. [15], the frequency of oscillations of the figure height gives the frequency of the gait. Positions of maxima in the height correspond to the positions in the gait where the swinging leg is vertical, thus defining a phase reference.

Conversion of Silhouettes to Features

A necessary step in silhouette-based gait recognition is conversion of a temporal sequence of silhouettes into a *gait signature*, i.e., a feature vector suitable for classification. One approach is to extract features that characterize the silhouette shapes and their variation over time, as illustrated schematically in Fig. 1a.

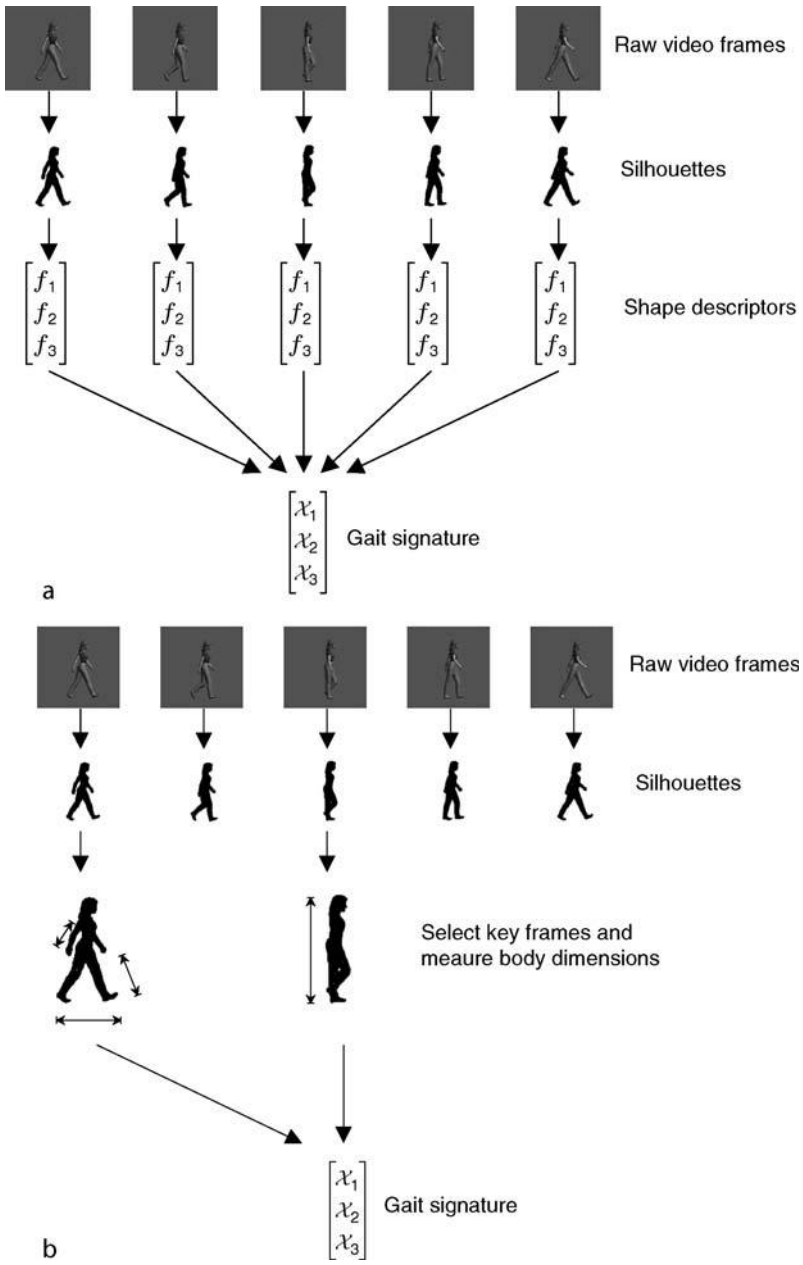
As an example, Little and Boyd [14] use geometric moments to describe a silhouette within a single frame of video. The moments include geometric centers, i.e., the average position of pixels in the silhouette, sometimes called the center of mass. Weighting the pixel positions by corresponding optical flow values gives geometric moments sensitive to rapid limb movement. Little and Boyd also use eccentricity [16], based on higher-order geometrical moments. A further step is necessary to combine the shape description for silhouettes in individual frames to a feature vector representative of the entire gait. Cyclic oscillations in the silhouette shape moments result naturally from a gait, so Little and Boyd exploit this to collect the individual shape descriptions into a single feature vector of the relative phases of the moment oscillations. Shutler and Nixon [17] describe a variation on this approach that uses Zernike moments to represent an accumulated shape over the duration of a gait cycle.

Ben-Abdelkader et al. [18] also exploit the periodic nature of a gait to form feature vectors. Periodicity and symmetry in a gait mean that similar shapes occur throughout the cycle of a gait. A feature vector built from measures of the silhouette self-similarity over period forms the basis for gait recognition. Periodicity in the self-similarity measures establishes the frequency of the gait. Hayfron-Acquah et al. [19] characterize the silhouette shape in a single frame by measuring symmetries in the outline of the silhouette to produce a *symmetry map*. The average of these symmetry maps over a gait cycle gives the gait signature used for recognition. Boyd [20] uses an array of phase-locked loops to measure the frequency, amplitude, and phase of pixel intensity oscillations due to a gait. The amplitudes and relative phases form a vector of complex phasors that acts as gait signature for recognition.

Rather than relying on the connection between gait and body structure to form a gait signature, one can use feature vectors that relate directly to body dimensions as shown in Fig. 1b. For example, Bobick and Johnson [21] measure stride and torso lengths, and Ben-Abdelkader et al. [15] measure height and stride characteristics. Collins et al. [22] identify key frames in a gait sequence for both the double-support (two feet on the ground) and mid-stride phase of a gait. From these key frames they measure cues related to height, width, and other body proportions, and movement-related characteristics such as stride length, and amount of arm swing.

Data Sets

A database of sample gaits is essential for developing a silhouette-based gait recognition system. Little and Boyd [14] provided one of the earliest databases

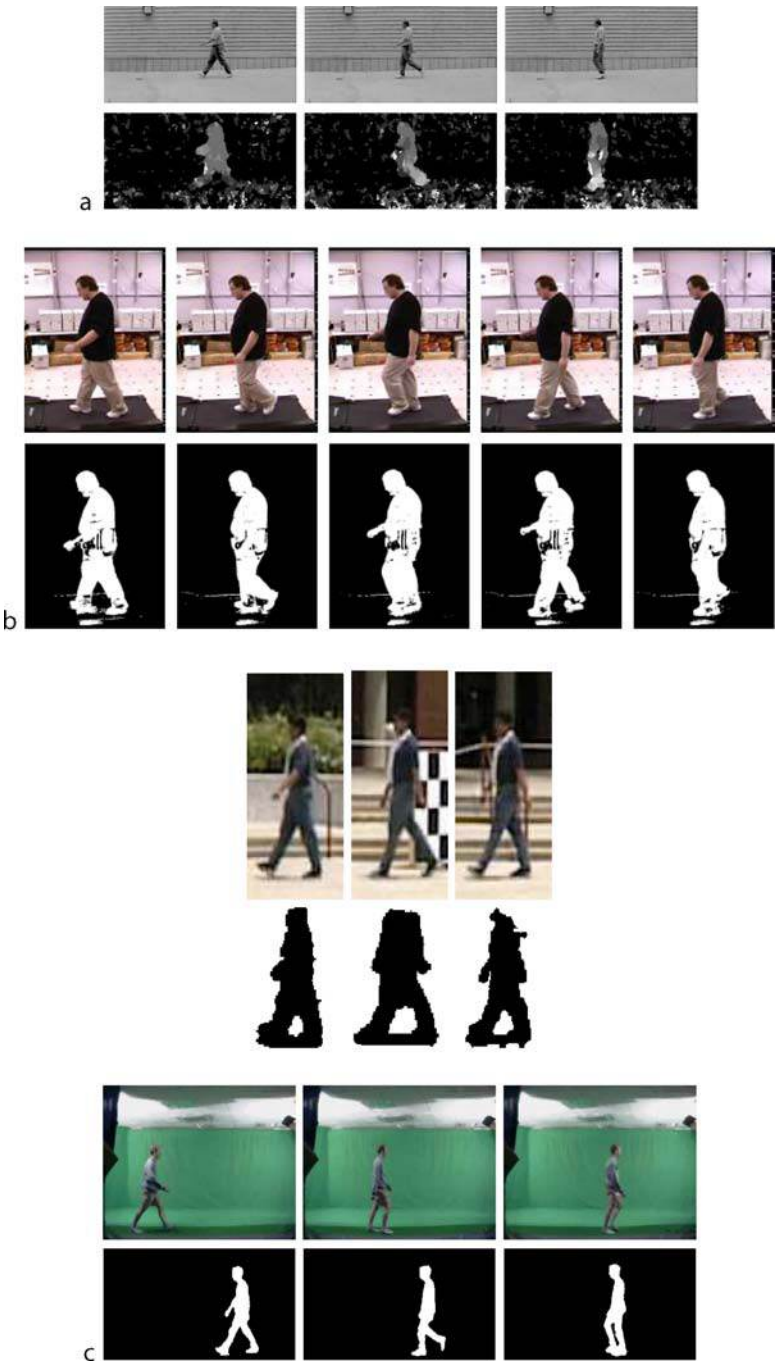


Gait Recognition, Silhouette-Based. **Figure 1** Themes in silhouette-based gait recognition: (**a**) shape descriptors of the silhouette combine to form a gait signature from the motion of the gait, or (**b**) critical body dimensions are measured from key frames within the gait cycle. Existing methods use variations on both of these themes and can even combine them.

featuring seven sample gaits for each of six subjects, for a total of 42 gait sequences (Fig. 2a).

Gross and Shi [23] created the Motion of Body (MOBO) database (Fig. 2b). It features gait samples for 25 subjects. Each subject walks on a treadmill under four different conditions (slow, fast, on an incline, and carrying a ball) and from a variety of viewing angles. Segmented silhouettes are part of the database.

Sarkar et al. [24] present a large (1.2 Gigabytes) gait database as part of the *HumanID Gait Challenge Problem* associated with the Defense Advanced Research Projects Agency (DARPA) HumanID project (Fig. 2c).



Gait Recognition, Silhouette-Based. **Figure 2** Example images from gait databases suitable for testing silhouette-based gait recognition: (**a**) Little and Boyd [14], (**b**) MOBO [21], (**c**) HumanID Gait Challenge [22], and (**d**) Shutler et al. [23]. All examples show raw video images in the top row and silhouettes or magnitude of the optical flow (Little and Boyd only) in the bottom row. The silhouettes shown for the Shutler et al. do not correspond to the images above.

The database contains samples for 122 subjects acquired in multiple sessions and under variable conditions. The *challenge problem* specifies a series of tests using the database as well as a reference algorithm to facilitate comparative testing by researchers.

Shutler et al. [25] created a database featuring over 100 subjects (Fig. 2d). The database contains sequences acquired over multiple sessions and features subjects walking from both left-to-right and right-to-left. Subjects walk on the ground or on treadmills, and

in front of green screens (for chroma-keying) or in outdoor scenes.

Examples

Bhanu and Han [26] estimate upper bounds on the performance of gait recognition by equating gait with body dimensions, presented as plots of recognition rate versus gallery size for varying assumptions of accuracy. As one might expect with upper bounds, these rates are optimistic. Random guessing is a good lower bound on performance, but any practical biometric system must be much better. One can reasonably expect that a gait biometric system should perform at least as well as humans on moving light displays [2], i.e., 38% from a gallery of six.

Within these broad bounds, there are numerous examples of existing silhouette-based gait recognition systems. Most of these have been tested with one or more of the databases mentioned earlier. Examples include the work of Hayfron-Acquah et al. [19], Shutler and Nixon [17], Collins et al. [22], Bobick and Johnson [21], Ben-Abdelkader et al. [15, 18], Liu and Sarkar [27], Robledo and Sarkar [28]. Lee and Grimson [29], Little and Boyd [14, 20], and Wang et al. [30]. The best reported correct classification rates (CCR) are better than 90% from a gallery of approximately 100 people.

Summary

Human experience supported by psychological observation suggests that humans can be recognized by their gaits, which inspires gait biometric systems. Silhouette-based gait recognition systems convert images from a video gait sequence to silhouettes of the walker. Dynamic shape or body dimensions are measured from the silhouettes and combined to form a gait signature used for recognition. There are several databases available for testing silhouette-based gait recognition, and numerous published examples of successful recognition using these databases.

Related Entries

- ▶ Gait Recognition, Model-Based
- ▶ Gait Recognition, Motion Analysis for
- ▶ Human Detection and Tracking

References

1. Boyd, J.E., Little, J.J.: Advanced Studies in Biometrics: Summer School on Biometrics, Alghero, Italy, June 2–6, 2003, *Revised Selected Lectures and Papers (Lecture Notes on Computer Science)*, vol. 3161/2005, chap. Biometric Gait Recognition, pp. 19–42. Springer (2005)
2. Cutting, J.E., Kozlowski, L.T.: Recognizing friends by their walk: gait perception without familiarity cues. Bull. Psychonomic Soc. **9**(5), 353–356 (1977)
3. McGeer, T.: Passive dynamic walking. Int. J. Robot. Res. **9**(2), 62–82 (1990)
4. McGeer, T.: Passive walking with knees. In: IEEE International Conference on Robotics and Automation, pp. 1640–1645 (1990)
5. Kuo, A.D.: A simple model of bipedal walking predicts the preferred speed-step length relationship. J. Biomech. Eng. **123**, 264–269 (2001)
6. Kuo, A.D.: Energetics of actively powered locomotion using the simplest walking model. J. Biomech. Eng. **124**, 113–120 (2002)
7. von Tscharner, V., Goepfert, B.: Gender dependent emgs of runners resolved by time/frequency and principal pattern analysis. J. Electromyogr. Kines. **13**, 253–272 (2003)
8. Baumberg, A.M., Hogg, D.C.: Generating spatiotemporal models from examples. In: Sixth British Conference on Machine Vision, Vol. 2, pp. 413–422. Birmingham, UK (1995)
9. Brinkmann, R.: The Art and Science of Digital Compositing. Morgan-Kaufmann, San Diego, CA (1999)
10. Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Computer Vision and Pattern Recognition, vol. II, pp. 246–252. Fort Collins, CO, USA (1998)
11. Barron, J.L., Fleet, D.J., Beauchemin, S.: Performance of optical flow techniques. Int. J. Comput. Vision **12**(1), 43–77 (1994)
12. Troje, N.F.: Decomposing biological motion: a framwork for analysis and synthesis of human gait patterns. J. Vision **2**, 371–387 (2002)
13. Omlor, L., Giese, M.A.: Extraction of spatio-temporal primitives of emotional body expressions. Neurocomputing **70**(10–12), 1938–1942 (2007)
14. Little, J.J., Boyd, J.E.: Recognizing people by their gait: the shape of motion. Videre **1**(2), 1–32 (1998)
15. Ben-Abdelkader, C., Cutler, R., Davis, L.: Person identification using automatic height and stride estimation. In: 16th International Conference on Pattern Recognition, pp. 377–380. Quebec, Quebec (2002)
16. Ballard, D.H., Brown, C.M.: Computer Vision. Prentice-Hall, Englewood Cliffs, NJ (1982)
17. Shulter, J.D., Nixon, M.S.: Zernike velocity moments for sequence-based descriptiion of moving features. Image and Vision Computing **24**, 343–356 (2006)
18. Ben-Abdelkader, C. Cutler, R., Davis, L.: Motion-based recognition of people in eigengait space. In: 5th IEEE International Conference on Automatic Face and Gesture Recognition, 267–272 (2002)
19. Hayfron-Acquah, J.B., Nixon, M.S., Carter, J.N.: Automatic gait recognition by symmetry analysis. Pattern Recogn. Lett. **24**, 2175–2183 (2003)
20. Boyd, J.E.: Synchronization of oscillations for machine perception of gaits. Comput. Vision Image Understand. **96**, 35–59 (2004)

21. Bobick, A., Johnson, A.: Gait recognitin using static activity-specific parameters. In: Computer Vision and Pattern Recognition 2001, Vol. I, pp. 423–430. Kauai, HI (2001)
22. Collins, R.T., Gross, R., Shi, J.: Silhouette-based human identification from body shape and gait. In: Automatic Face and Gesture Recognition, pp. 351–356. Washington DC (2002)
23. Gross, R., Shi, J.: The cmu motion of body (mobo) database. Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University (2001)
24. Sarkar, S., Phillips, J., Liu, Z., Robledo, I., Prother, P., Bowyer, K. W.: The humanid gait challenge problem: data sets, performance, and analysis. IEEE Trans. Pattern Analy. Mach. Intell. **27**(2), 162–177 (2005)
25. Shutler, J., Grant, M., Nixon, M., Carter, J.: On a large sequence-based human gait database. In: Fourth International Conference on Recent Advances in Soft Computing, pp. 66–71. Nottingham, UK (2002)
26. Bhanu, B., Han, J.: Bayesian-based performance prediction for gait recognition. In: IEEE Workshop on Motion and Video Computing, pp. 145–150. Orlando, Florida (2002)
27. Liu, Z., Sarkar, S.: Improved gait recognition by gait dynamics normalization. IEEE Trans. Pattern Analy. Mach. Intell. **28**(6), 863–876 (2006)
28. Robledo, I., Sarkar, S.: Statistical motion model based on the change of feature relationships: human gait-based recognition. IEEE Trans. Pattern Analy. Mach. Intell. **25**(10), 1323–1328 (2003)
29. Lee, L., Grimson, W.: Gait analysis for recognition and classification. In: Automatic Face and Gesture Recognition, pp. 148–155. Washington DC (2002)
30. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. IEEE Trans. Pattern Analy. Mach. Intell. **25**(12), 1505–1518 (2003)

Gait, Forensic Evidence of

Niels Lynnerup, Peter K. Larsen
Laboratory of Biological Anthropology, Institute of Forensic Medicine, University of Copenhagen, Copenhagen, Denmark

Synonyms

Gait analysis; Perpetrator identification

Definition

Forensic evidence of gait, or forensic gait analysis, may be defined as analyses of gait performed in the service of the law. Usually, this involves analyses of criminal cases with the aim to characterize the gait of a perpetrator, and often to compare the gait of a perpetrator with the gait of a suspect. The results of the analyses may furthermore need to be presented in court. The methods involved in forensic gait analyses comprise morphological assessment of single gait features and kinematic assessment of body movement, often combined with photogrammetrics. The latter means that body segment lengths, stride length, etc. may be quantified and used in direct comparisons.

Introduction

Forensic analysis of ▶ gait has a lot of common ground with biometric gait recognition, but there are also some major differences. In terms of image capture, the imagery used in forensic gait analysis is mostly always acquired from CCTV, with the perpetrator specifically trying to conceal identity. In biometric systems, image capture of a person, or registration, takes place under specific circumstances, designed to maximize data quality, and obviously a person will willingly follow a set of guidelines in order to ensure proper registration. On the other hand a bank robber might try to hide his or her face to avoid facial recognition or wear baggy clothes to blur body morphology.

Biometric gait recognition systems may operate with various false accept or reject rates, which govern how exclusive the system is, and reflect the number of "wrong" registrations that can be tolerated. For example, a relatively high false reject rate (i.e., rejecting a person who otherwise should be cleared) is not a problem if the system is meant for a screening function, where rejection simply leads to an additional identity check. It is possible to generate computer models which can identify people by their gait with more than 90% success [1, 2], but these models are still based on a small number of people and require optimal conditions seldom found outside the laboratory [3]. Alternative biometric approaches use a description of a subject's silhouette, often with reportedly improved recognition performance [4]. In forensic gait analysis, the analysis is often specifically carried out to match a perpetrator with a suspect. If the case is made that there is a match then the suspect may be sentenced. This places a certain onus on the gait analysis and the scientists carrying out the analyses, and the prosecution and the defense may well challenge the findings of the

gait analysis. This also means that when presenting the results of a forensic gait analysis, one has to be familiar with the legal prerequisites for the legal statements, and how expert evidence is adjudicated.

Technical Issues

Bank CCTV systems are often set up not to capture gait specifics, but rather to give fields of view covering office spaces, teller machines, etc., and also often to supervise the bank employees. It is a not uncommon experience when perusing CCTV footage after a bank robbery that the perpetrator is seen moving behind desks and tellers, so that only the upper part of his body is filmed. Also, the CCTV system may vary quite a lot in terms of technical quality, e.g., image capture frequencies, digital versus analog data storage, color versus b/w cameras (the latter often of sharper quality), and numerous supplier – dependent video and computer systems (e.g., in terms of data compression of video images).

The recording frequency should ideally be about 15 Hz allowing the examination of dynamic features such as, e.g., lateral instability in the knee at heel strike. Others have found a similar frequency sufficient for obtaining joint angles [5] and for automatic recognition of gait [2]. Lower recording frequencies may also be sufficient to examine features that are more static, although the gait will have a "jerky" appearance. Even at a low 5 Hz recording frequency, it has proved possible to examine gait parameters such as dorsal/plantar flexion at heel strike, degree of "push-off" at toe-off, and knee flexion during stance. At even lower recording frequencies, where the images really are still image series, specific gait-related characteristics may be noticed, e.g., a perpetrator with a bow-legged left knee. This means that even just one single image of the gait can sometimes be useful, if the gait feature captured can be deemed characteristic.

Gait

The ability to recognize other individuals is fundamental to human life. Identification by gait is a part of this process. Shakespeare made use of this in his play "The Tempest" where Ceres said: "High'st queen of state, Great Juno, comes; I know her by her gait". Psychophysiological studies have proved that the human being can recognize the sex of a walker [6] and friends and colleagues [7, 8] with a success rate up to 70–80%.

The authors derive from the Institution that has conducted what are, so far, the only scientific approaches to gait analysis for evidential procedures. The essay describes how evidential analysis was derived and presented in two forensic investigations [9, 10].

Gait analyses is performed by first gaining a purely morphological, ▶ anthroposcopic impression of the gait of a perpetrator. We then combine the basic ability to recognize people with biomechanical knowledge in order to give statements as to whether a suspect could have the same identity as a perpetrator in a given case by comparing the suspect's posture and joint angles during gait with the perpetrator's. A checklist has been developed for forensic gait analysis (Table 1). First described are the general characteristics of the perpetrator's gait following which are analyzed each of the joint rotations and segment movements found relevant for forensic gait analysis (by trial end error). When a profile of the perpetrator has been completed, each item of the list is compared to the recording of the suspect and stated if agreement (A), no agreement (N), or comparison not possible (-) is found. An item can be incomparable because either the joint rotation/movement cannot be analyzed due to poor quality of the surveillance recordings, or the recording of the suspect differs too much in some way from the recording of the crime such as differences in shoulder angles between suspect and perpetrator because of elevated shoulders in one of the recordings.

There have been several automated assessments of feature analysis for forensic and biometric purposes which show that there is a natural match between technique and observed performance [5]. Their features include foot angle (degree of outward rotation), the step length, and the mean hip joint angle, among others. Several other characteristic features have also been identified: inversion/eversion in the ankle during stance, lateral flexion in the dorsal column of the spine, and the knee angle in the frontal plane that would show lateral instability of the knee and signs of a person being bow-legged/knock-kneed. Furthermore, some of the characteristic features found were so special, such as limping, that it was not necessarily expected to be found in the 11 randomly selected subjects.

Gait, Forensic Evidence of. Table 1 IFM Copenhagen gait description form/checklist. The rightmost column is marked up either with "A" for agreement; "N" for no agreement; and "-" for incomparable (see text).
The middle column is used for notes and specific observations

General	Notes on gait of perpetrator/ suspect	
Long/short steps, stiff/ relaxed gait with Narrow/ wide distance between the feet		
Signs of pathologic gait		
Feet/ankle joint		
Outward rotation		
Inversion/eversion		
Dorsal/plantar flexion at heel strike		
Degree of "push-off" at toe-off		
Knee		
Varus/valgus		
Knee flexion during stance		
Hip/pelvis		
Pelvis Abduction/adduction		
Pelvis Rotation		
Pelvis tilt		
Upper body		
Lateral flexion of spinal column		
Forward/backward leaning		
Rotation of the upper body during walk		
Shoulders		
Angle in frontal plane		
Forward/backward rotation		
Neck/head		
Posture in sagittal plane		
Head movements in frontal plane		
Quality of recordings/ other precautions		

It should be stressed that a rather wide definition of "gait analysis" is used, so that basically all bodily movements may be studied. Posture and stance may be quite specific. For example, when standing, one leg is more often weight-bearing than the other; there may be marked lordosis; the neck and shoulders may be more or less slouched, and so on. These stance-related characteristics have a bearing on how a person initiates or stops walking, and should thus also be involved in the analysis.

All the above features may be judged purely morphologically, but it may be of great evidentiary value to attach numbers to these features. Thus, the morphological approach is combined with photogrammetry in order to acquire specific measurements of body segment lengths and heights.

Photogrammetry in Association with Gait Analysis

Photogrammetry literally means measuring by photography. Photogrammetry enables the measurement of unknown values in two-dimensional space (2D) using known values within a single image [9, 10]. Another basic application of photogrammetry is measuring objects in three-dimensional space (3D) using photographs taken from different sides and angles. Zhao et al. [11] have also worked with video sequences in this respect. Jensen and Rudin [9] used a 2D method to measure the stature and several segment lengths in two different cases and found excellent agreement between perpetrator and suspect. Lynnerup and Vedel [10, 12] used a 3D method in the investigation of a bank robbery where the perpetrator was recorded simultaneously from two different cameras and found good agreement in bodily measurements when comparing the perpetrator to the suspect.

A first step in photogrammetry is calibration of the CCTV cameras. This is done by placing frames with targets on the locations (Fig. 1). The frames are photographed with both the surveillance video cameras and a calibrated digital camera. Using the digital camera images and special software (PhotoModeler Pro®) the points are measured and subsequently imported as control points ("fiduciary points"). A feature in PhotoModeler Pro® allows determination of the internal parameters of the surveillance video cameras, e.g., focal length, and subsequent calculation of the exact placement of the cameras. After calibration, still images from the surveillance cameras are input in PhotoModeler Pro®. The photogrammetrical method described here has the advantage that there is no need

to ascertain the position of the perpetrator in relation to a measuring device. After calibration by fiduciary points, the photogrammetrical analysis produces points in a 3D space, and an evaluation of the goodness of fit may be made directly in the software. This then allows measuring body segment lengths, stature, etc. of a perpetrator in various locations and with various body stances (Fig. 2). The selection of anatomical points is done by choosing specific points such as the top of the head, eyes, and joint center-points on an image. This selection is made by judging anatomical landmarks, clothing displacement, comparison with images just before and after the chosen photo, etc. When then focusing on the other images of the same situation, but from other cameras, the program will indicate the epi-lines (the "line of sight") from the first image, as well as a line connecting the two joints. After selecting the identical anatomical points in this image, it is immediately apparent how good the fit is, and whether the points selected in the first image are adequate. Thus, the 3D coordinates are calculated not only by a simple averaging of points chosen from two images, but reflect a dynamic process where the tightness of the intersections of the epi-lines is minimized. The absolute error associated with measuring using photogrammetry as described is small. For instance, the height of a desk (bolted to the floor and not moved between the incident and the analysis) was measured by photogrammetry (result: 89.3 cm) and compared to

Gait, Forensic Evidence of. **Figure 1** Measuring screens put up in a department store, in order to calibrate the CCTVs [10].

Gait, Forensic Evidence of. **Figure 2** Screen shots of PhotoModeler Pro® interface, showing selection of points. Simultaneous images from different CCTV cameras are used to pinpoint concurrent anatomical points (and markers) seen from different POV. The lines between the points are to scale and thus hold accurate measures of distance [10].

an actual physical measurement (result: 90.0 cm), thus the error was 7 mm or less than 1%. Intra- and inter-observer tests of photogrammetric measurements of bodily segments seem to indicate that the error associated with clearly identifiable body points, such as top of the head, eyes, ear lobes, among others is small. On the other hand, if the body points are hidden or obscured by clothing, such as joint center-points, then there is some variation, which needs to be taken into account.

Currently, research focuses on implementing the possibility of performing accurate measurements of a perpetrator even though images are from only one camera. To do this, a measuring screen is used, the contours of which can be accurately measured by the software, which is physically placed near to where the perpetrator was standing (it needs to place the perpetrator on a specific point on the floor). If the screen is oriented perpendicular to the camera, then the screen can be imported as a virtual screen overlaid the crime video-footage (Fig. 3). The perpetrator can then be measured against this screen, akin to seeing a person standing in front of a light-source, and whose shadow is cast of a screen or wall behind him.

Comparing Gait and Photogrammetry

As the forensic analysis mostly pertains to comparisons of perpetrators and suspects, then gait analyses and measuring of the suspect also has to be carried out. Owing to legal exigencies, this may be performed under very different settings and conditions, comprising hidden and overt image capture for gait, and hidden and overt photogrammetry. In some cases, legal circumstances have ruled out hidden image capture; in other cases the defense counsel was invited to be present (but without the knowledge of the suspect); and finally, the suspect has sometimes been filmed completely overt. Ideally, it is felt that gait image capture should be performed hidden, so as the suspect does not know he is being filmed. This is to ensure that the gait is not "changed". Preferably, the setting for performing the image capture should to some extent mimic the crime scene. For example, if at the crime scene there was a step at the entrance, which the suspect engaged in a distinct fashion, then filming the suspect engaging a somewhat likewise step would be obvious for comparison. If the crime scene images show a perpetrator walking down a corridor, either against or away from the CCTV camera, then a setting at police offices with a long corridor may be suitable. The filming usually takes place with ordinary DV-cameras, and is done by forensic technicians, but the setting would have been discussed in advance. For instance, a policeman can be instructed to accompany the suspect, but walking at a speed that matches the velocity of the perpetrator, because the gait speed may influence some of the features. For example, a lateral instability in the knee will be more pronounced at a higher gait speed.

Gait, Forensic Evidence of. Figure 3 Using the back-projection screen method (see text).

The photogrammetric measurement of the suspect is most easily performed overtly. Usually a corner in an office is identified with points fixed on the wall, and the suspect is asked to stand in the corner. Using two or three digital cameras, coupled to a computer, several sets of images of high quality for subsequent photogrammetry can be rapidly acquired. While height could be just as easily acquired using a stadiometer, it is found that the same measuring method (photogrammetry) should be used for comparing perpetrator and suspect. While at first glance stadiometer-measured stature might seem as a "gold-standard", it is also found that people almost automatically straighten themselves when asked to stand against a stadiometer, meaning in fact that a better agreement between subsequent measurings of stature by photogrammetry has been found, than between photogrammetry and a stadiometer. Of course, measuring the suspect by photogrammetry also makes it easier to measure other heights, such a floor to eye, floor to shoulder, and floor to ear-lobe.

Schöllhorn et al. [13] concluded that "identification of individuality seems to be impossible with single variables or specific parameters of single variables", so the more gait characteristics and bodily measurements of the perpetrator that can be extracted and compared to the suspect, the better.

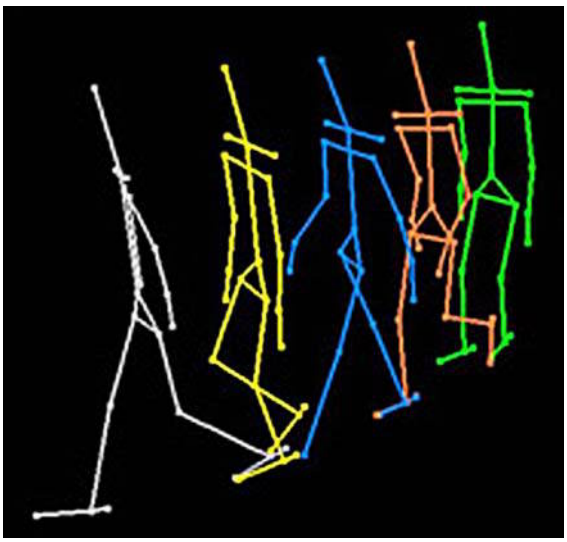
The Nature of Forensic Statements

In statements to the police it is noted what image material has been available, and what manner of image enhancing techniques had been used. The results of the above analyses are then presented, each followed by a separate conclusion, and each conclusion always summing up what features were found to indicate concordance between the suspect and the perpetrator, as well as features which seemed to indicate incongruity. Each item may therefore be seen as constituting single pieces of evidence. This renders a statistical approach, for instance the calculation of likelihood ratios for identity, based on the prevalence of certain facial and bodily traits, problematical [14].

Using the data sheets for gait analyses and photogrammetry fulfils three of the four guidelines in the ▶ Daubert Standard, a legal precedent set by Supreme Court of the United States [15], for determining whether expert witnesses' testimony is admissible as evidence: (1) the testimony in court is based on an empirically used technique, (2) the technique has been published in peer-reviewed literature and (3) it is generally accepted for use in forensic medicine. The last Daubert Guideline states that the reliability of the technique has been tested and potential error rates known.

Image based comparison will probably never achieve specific identification such as associated with DNA-typing and fingerprinting. However, analyzing gait and measuring stature and segment lengths of a perpetrator from surveillance video has the possibility of becoming a valuable forensic tool because the gait and the measures are an integrated part of the offender. At present, the methods can be used effectively to exclude a suspect if the gait and anthropometrical measures of the suspect and perpetrator are entirely different from each other. On the other hand, if the perpetrator and suspect do have a similar gait and similar measures, it can only be stated in court that the suspect cannot be excluded as the perpetrator. To give a more specific statement of the value of evidence, a database with gait characteristics and measures for a population of which the perpetrator and suspect could be referenced against. In theory, this might mean that if a perpetrator and a suspect are measured to have an unusual height, i.e., either very tall or very small, then this might in itself increase the likelihood of concordance between them, whereas very average heights would lower the likelihood (because then it might be almost anybody). In actuality, such databases are rather restricted, with often only specific subsamples of the entire population represented; populations also change in terms of e.g., immigration; and finally the perpetrator might well be from an entirely different part of the world. If comparing with such databases, it is important to stress that "given the perpetrator/suspect are drawn from the same population as the database", then their stature is more or less common, and the likelihood of concordance between them is more or less likely.

Future work will probably focus on a better integration of gait characteristics and photogrammetry in order to perform dynamic measurements of gait (basically "animating" the line models, cf. Fig. 4). This has the potential of calculating angles of flexion – extension in the major joints, step length, degree of side-to-side movement of the torso during walking, etc. These parameters may then further assist in discriminating between suspects and more specifically in identifying individual traits of gait.



Gait, Forensic Evidence of. Figure 4 Line models produced by the Photomodeler® based on the selected anatomical points, showing the gait.

Related Entries

- ▶ Gait Recognition, Model-Based
- ▶ Gait Recognition, Motion Analysis for
- ▶ Gait Recognition, Overview
- ▶ Gait Recognition, Silhouette-Based

References

1. Cunado, D., Nixon, M.S., Carter, J.N.: Automatic extraction and description of human gait models for recognition purposes. Comput. Vis. Image Underst. **90**, 1–41 (2003)
2. Urtasun, R., Fua, P.: 3D tracking for Gait Characterization and Recognition. 1*Computer Vision Laboratory.* Swiss Federal Institute of Technology, Lausanne, Switzerland. Report No. IC/2004/04(2004)
3. Nixon, M.S., Tan, T.N., Chellappa, R.: Human identification based on gait. Springer Science + Business Media, Inc. New York, NY, USA (2006)
4. Rahati, S., Moravejian, R., Kazemi, F.M.: Gait Recognition Using Wavelet Transform Information Technology: New Generations, 2008. Fifth International Conference on Information Technology. pp. 932–936. 7–9, proceedings, Universitat Triet, Germany (April 2008)
5. Wagg, D.K., Nixon, M.S.: On automated model-based extraction and analysis of gait, Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, IEEE Seoul, Korea. 11–19. 17–19 (May 2004)
6. Kozlowski, L.T., Cutting, J.E.: Recognizing the sex of a walker from a dynamic point-light display. Percept. Psychophys. **21**, 575–580 (1977)
7. Cutting, J.E., Kozlowski, L.T.: Recognizing friends by their walk: Gait perception without familiarity cues. Bull. Psychon. Soc. **9**, 353–356 (1977)
8. Jokisch, D., Daum, I., Troje, N.F.: Self recognition versus recognition of others by biological motion: viewpoint-dependent effects. Perception **35**(7), 911–920 (2006)
9. Jensen, S.C., Rudin, L.I.: Measure: an interactive tool for accurate forensic photo/videogrammetry. In: Rudin, L.I., Bramble, S.K. (eds.) Investigative and Trial Image Processing 1995 July 13; 73–83. SPIE, San Diego, CA (1995)
10. Lynnerup, N., Vedel, J.: Person identification by gait analysis and photogrammetry. J. Foren. Sci. **50**(1), 112–118 (2005)
11. Zhao, G., Liu, G., Hua, L., Pietikainen, M.: 3D gait recognition using multiple cameras. In Automatic Face and Gesture Recognition, 2006. FGR 2006. Seventh International Conference, 10–12 April 2006: 529–534 (2006)
12. Lynnerup, N., Sejrsen, B., Vedel, J.: Identification by facial recognition, gait analysis and photogrammetry: The Anna Lindh murder. In: Brickley, M., Ferllini, R. (eds.) Forensic Anthropology: Case studies from Europe, pp. 232–244, Charles C. Thomas, Springfield, Ill., U.S.A. (2007)
13. Aitken, C.G.G.: Statistics and the evaluation of evidence for forensic scientists. Wiley and Sous, Chichester, UK (1995)
14. *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993)
15. Schollhorn, W.I., Nigg, B.M., Stefanyshyn, D.J., Liu, W.: Identification of individual walking patterns using time discrete and time continuous data sets. Gait Posture **15**(2), 180–186 (2002)

Gallery and Probe

Gallery is one of the data partitions in an algorithm-level biometric evaluation experiment. It is a collection of biometric templates that form the search dataset. Typically, these are representative of the enrolled templates in an actual biometric deployment scenario. In algorithm-level evaluations, care should be taken to have same number of representative templates per subject in the gallery. Probe is the second data partition in an algorithm-level evaluation experiment. It is a collection of biometric templates that need to be recognized or identified by matching against the gallery. In any given algorithm-level evaluation, the probes and gallery differ with respect to the covariate that is being studied. For example, to study the impact of viewpoint covariate, the gallery is chosen to be from one viewpoint and the probe is chosen to be from a different viewpoint. Since during actual operations biometric data is expected to arrive in a sequential fashion, it is not appropriate to normalize or adjust biometric

matching scores over the probes. Neither is it appropriate to train on the probe data.

► Evaluation of Gait Recognition

Gaussian Mixture Density

► Gaussian Mixture Models

Gaussian Mixture Models

Douglas Reynolds
Lincoln Laboratory, MIT, Lexington, MA, USA

Synonyms

Gaussian mixture density; GMM

Definition

A Gaussian Mixture Model (GMM) is a parametric ► probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as vocal-tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or ► Maximum *A Posteriori* (MAP) estimation from a well-trained prior model.

Introduction

A Gaussian mixture model is a weighted sum of M component Gaussian densities as given by the equation,

$$p(\mathbf{x}|\lambda) = \sum_{i=1}^{M} w_i \, g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \qquad (1)$$

where $\mathbf{x}$ is a D-dimensional continuous-valued data vector (i.e. measurement or features), w_i, $i = 1, \ldots, M$, are the mixture weights, and $g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, \ldots, M$ are the component Gaussian densities. Each component density is a D-variate Gaussian function of the form,

$$g(\mathbf{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)' \, \Sigma_i^{-1} \, (\mathbf{x}-\boldsymbol{\mu}_i)\right\}, \qquad (2)$$

with mean vector $\boldsymbol{\mu}_i$ and covariance matrix Σ_i. The mixture weights satisfy the constraint that $\sum_{i=1}^{M} w_i = 1$.

The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation,

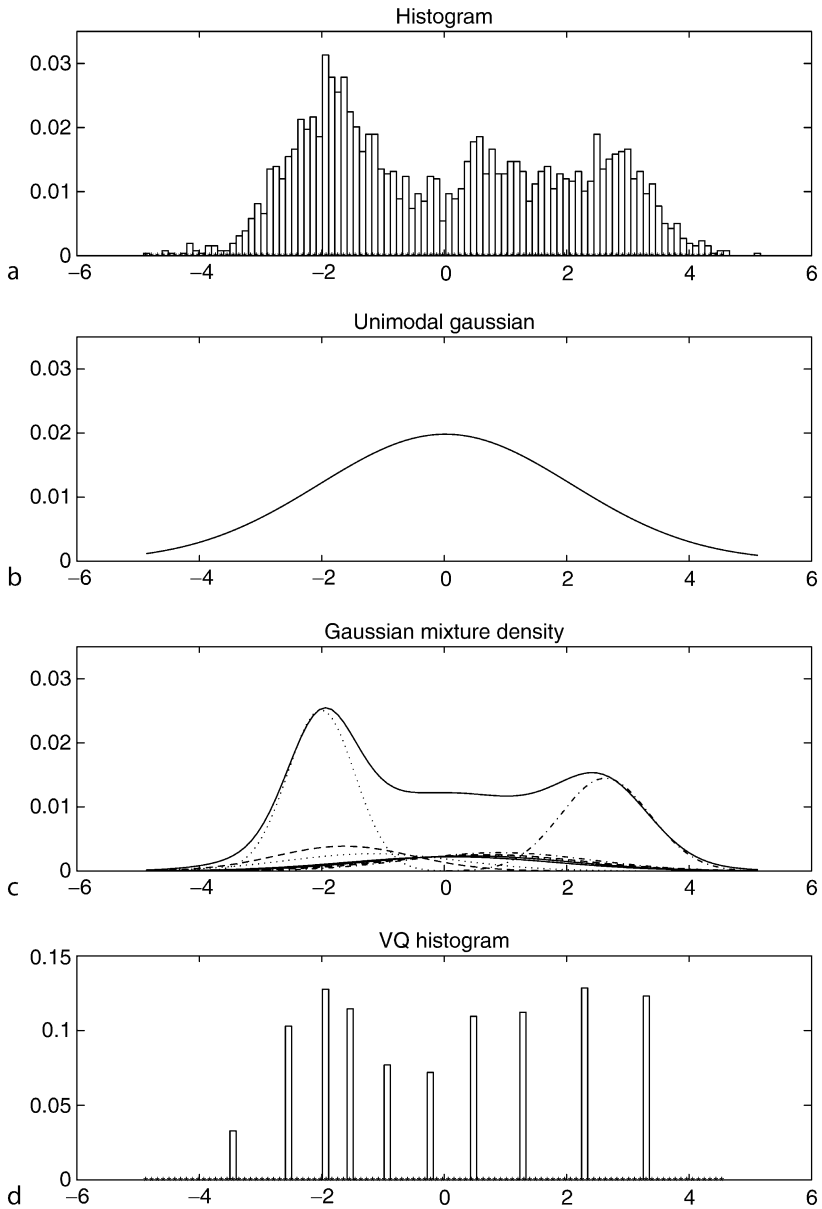
$$\lambda = \{w_i, \boldsymbol{\mu}_i, \Sigma_i\} \qquad i = 1, \ldots, M. \qquad (3)$$

There are several variants on the GMM shown in Eq. (3). The covariance matrices, Σ_i, can be full rank or constrained to be diagonal. Additionally, parameters can be shared, or tied, among the Gaussian components, such as having a common covariance matrix for all components, The choice of model configuration (number of components, full or diagonal covariance matrices, and parameter tying) is often determined by the amount of data available for estimating the GMM parameters and how the GMM is used in a particular biometric application.

It is also important to note that since the component Gaussian are acting together to model the overall feature density, full covariance matrices are not necessary even if the features are not statistically independent. The linear combination of diagonal covariance basis Gaussians is capable of modeling the correlations between feature vector elements. The effect of using a set of M full covariance matrix Gaussians can be equally obtained by using a larger set of diagonal covariance Gaussians.

GMMs are often used in biometric systems, most notably in speaker recognition systems [1, 2], due to their capability of representing a large class of sample distributions. One of the powerful attributes of the GMM is its ability to form smooth approximations to arbitrarily shaped densities. The classical unimodal Gaussian model represents feature distributions by a position (mean vector) and a elliptic shape (covariance matrix) and a vector quantizer (VQ) or nearest neighbor model represents a distribution by a discrete set

of characteristic templates [3]. A GMM acts as a hybrid between these two models by using a discrete set of Gaussian functions, each with its own mean and covariance matrix, to allow a better modeling capability. Figure 1 compares the densities obtained using a unimodal Gaussian model, a GMM, and a VQ model. Plot (a) shows the histogram of a single feature from a speaker recognition system (a single cepstral value from a 25 second utterance by a male speaker); plot (b) shows a unimodal Gaussian model of this feature distribution; plot (c) shows a GMM and its ten underlying component densities; and plot (d) shows a histogram of the data assigned to the VQ centroid locations of a ten element codebook.



Gaussian Mixture Models. Figure 1 Comparison of distribution modeling. (**a**) histogram of a single cepstral coefficient from a 25 second utterance by a male speaker (**b**) maximum likelihood unimodal Gaussian model (**c**) GMM and its ten underlying component densities (**d**) histogram of the data assigned to the VQ centroid locations of a ten element codebook.

The GMM not only provides a smooth overall distribution fit, its components also clearly detail the multimodal nature of the density.

The use of a GMM for representing feature distributions in a biometric system may also be motivated by the intuitive notion that the individual component densities may model some underlying set of *hidden* classes. For example, in speaker recognition, it is reasonable to assume the acoustic space of spectral related features corresponding to a speaker's broad phonetic events, such as vowels, nasals, or fricatives. These acoustic classes reflect some general speaker-dependent vocal tract configurations that are useful for characterizing speaker identity. The spectral shape of the ith acoustic class can in turn be represented by the mean $\boldsymbol{\mu}_i$ of the ith component density, and variations of the average spectral shape can be represented by the covariance matrix Σ_i. Since all the features used to train the GMM are unlabeled, the acoustic classes are hidden in that the class of an observation is unknown. A GMM can also be viewed as a single-state HMM with a Gaussian mixture observation density, or an ergodic Gaussian observation HMM with fixed, equal transition probabilities. Assuming independent feature vectors, the observation density of feature vectors drawn from these hidden acoustic classes is a Gaussian mixture [4, 5].

Maximum Likelihood Parameter Estimation

Given training vectors and a GMM configuration, the parameters, λ, are estimated which, in some sense, best match the distribution of the training feature vectors. There are several techniques available for estimating the parameters of a GMM [6]. By far the most popular and well-established method is ▶ maximum likelihood (ML) estimation.

The aim of ML estimation is to find the model parameters which maximize the likelihood of the GMM given the training data. For a sequence of T training vectors $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_T\}$, the GMM likelihood, assuming independence between the vectors (The independence assumption is often incorrect but is needed to make the problem tractable.), can be written as,

$$p(X|\lambda) = \prod_{t=1}^{T} p(\mathbf{x}_t|\lambda). \tag{4}$$

Unfortunately, this expression is a nonlinear function of the parameters λ and direct maximization is not possible. However, ML parameter estimates can be obtained iteratively using a special case of the expectation-maximization (EM) algorithm [7].

The basic idea of the EM algorithm is, beginning with an initial model λ, to estimate a new model $\bar{\lambda}$, such that $p(X|\bar{\lambda}) \geq p(X|\lambda)$. The new model then becomes the initial model for the next iteration and the process is repeated until some convergence threshold is reached. The initial model is typically derived by using some form of binary VQ estimation.

On each EM iteration, the following re-estimation formulas are used which guarantee a monotonic increase in the model's likelihood value,

Mixture Weights

$$\bar{w}_i = \frac{1}{T} \sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda). \tag{5}$$

Means

$$\bar{\boldsymbol{\mu}}_i = \frac{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)\ \mathbf{x}_t}{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)}. \tag{6}$$

Variances (diagonal covariance)

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)\ \mathbf{x}_t^2}{\sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda)} - \bar{\mu}_i^2, \tag{7}$$

where ${\sigma_i}^2$, x_t, and μ_i refer to arbitrary elements of the vectors $\boldsymbol{\sigma}_i^2$, $\mathbf{x}_t$, and $\boldsymbol{\mu}_i$, respectively.

The *a posteriori* probability for component i is given by

$$\Pr(i|\mathbf{x}_t, \lambda) = \frac{w_i\ g(\mathbf{x}_t|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{\sum_{k=1}^{M} w_k\ g(\mathbf{x}_t|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)} \tag{8}$$

Maximum *A Posteriori* (MAP) Parameter Estimation

In addition to estimating GMM parameters via the EM algorithm, the parameters may also be estimated using Maximum *A Posteriori* (MAP) estimation. MAP

estimation is used, for example, in speaker recognition applications to derive speaker model by adapting from a universal background model (UBM) [8]. It is also used in other pattern recognition tasks where limited labeled training data is used to adapt a prior, general model.

Like the EM algorithm, the MAP estimation is a two-step estimation process. The first step is identical to the "Expectation" step of the EM algorithm, where estimates of the sufficient statistics (These are the basic statistics needed to be estimated to compute the desired parameters. For a GMM mixture, these are the count, and the first and second moments required to compute the mixture weight, mean and variance.) of the training data are computed for each mixture in the prior model. Unlike the second step of the EM algorithm, for adaptation these "new" sufficient statistic estimates are then combined with the "old" sufficient statistics from the prior mixture parameters using a data-dependent mixing coefficient. The data-dependent mixing coefficient is designed such that mixtures with high counts of new data rely more on the new sufficient statistics for final parameter estimation and mixtures with low counts of new data rely more on the old sufficient statistics for final parameter estimation.

The specifics of the adaptation are as follows. Given a prior model and training vectors from the desired class, $X = \{\mathbf{x}_1 \ldots, \mathbf{x}_T\}$, the probabilistic alignment of the training vectors into the prior mixture components is determined (Fig. 2a). That is, for mixture i in the prior model, $\Pr(i|\mathbf{x}_t, \lambda_{\text{prior}})$ is computed as in Eq. (8).

Then compute the sufficient statistics for the weight, mean, and variance parameters $\mathbf{x}^2$ is shorthand for $\text{diag}(\mathbf{x}\mathbf{x}')$:

$$n_i = \sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda_{\text{prior}}) \text{ weight}, \tag{9}$$

$$E_i(x) = \frac{1}{n_i} \sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda_{\text{prior}}) \mathbf{x}_t \text{ mean}, \tag{10}$$

$$E_i(\mathbf{x}^2) = \frac{1}{n_i} \sum_{t=1}^{T} \Pr(i|\mathbf{x}_t, \lambda_{\text{prior}}) \mathbf{x}_t^2 \text{ variance}. \tag{11}$$

This is the same as the "Expectation" step in the EM algorithm.

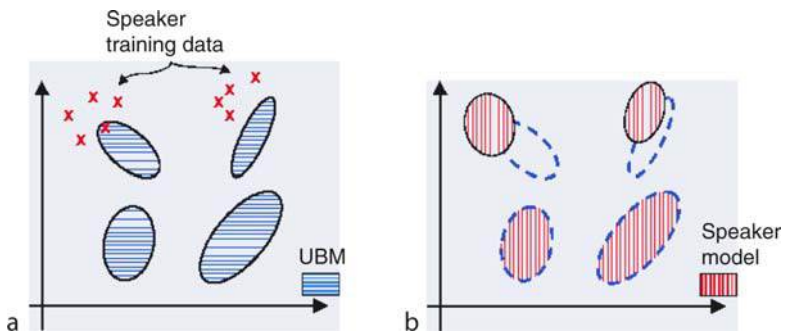
Lastly, these new sufficient statistics from the training data are used to update the prior sufficient statistics for mixture i to create the adapted parameters for mixture i (Fig. 2b) with the equations:

$$\hat{w}_i = \left[\alpha_i^w n_i / T + (1 - \alpha_i^w) w_i\right] \gamma \quad \text{adapted mixture weight}, \tag{12}$$

$$\hat{\boldsymbol{\mu}}_i = \alpha_i^m E_i(\mathrm{x}) + (1 - \alpha_i^m) \boldsymbol{\mu}_i \quad \text{adapted mixture mean}, \tag{13}$$

$$\hat{\boldsymbol{\sigma}}_i^2 = \alpha_i^v E_i(\mathbf{x}^2) + (1 - \alpha_i^v)(\boldsymbol{\sigma}_i^2 + \boldsymbol{\mu}_i^2) - \hat{\boldsymbol{\mu}}_i^2 \quad \text{adapted mixture variance} \tag{14}$$

The adaptation coefficients controlling the balance between old and new estimates are $\{\alpha_i^w, \alpha_i^m, \alpha_i^v\}$ for the weights, means, and variances, respectively. The scale factor, γ, is computed over all adapted mixture weights to ensure that they sum to unity. Note that the



Gaussian Mixture Models. **Figure 2** Pictorial example of two steps in adapting a hypothesized speaker model. (**a**) The training vectors (x's) are probabilistically mapped into the UBM (prior) mixtures. (**b**) The adapted mixture parameters are derived using the statistics of the new data and the UBM (prior) mixture parameters. The adaptation is data-dependent, so UBM (prior) mixture parameters are adapted by different amounts.

sufficient statistics, not the derived parameters, such as the variance, are being adapted.

For each mixture and each parameter, a data-dependent adaptation coefficient α_i^ρ, $\rho \in \{w, m, v\}$, is used in the equations mentioned earlier. This is defined as

$$\alpha_i^\rho = \frac{n_i}{n_i + r^\rho}, \tag{15}$$

where r^ρ is a fixed "relevance" factor for parameter ρ. It is common in speaker recognition applications to use one adaptation coefficient for all parameters ($\alpha_i^w = \alpha_i^m = \alpha_i^v = n_i/(n_i + r)$) and further to only adapt certain GMM parameters, such as only the mean vectors.

Using a data-dependent adaptation coefficient allows mixture-dependent adaptation of parameters. If a mixture component has a low probabilistic count, n_i, of new data, then $\alpha_i^\rho \to 0$ causing the de-emphasis of the new (potentially under-trained) parameters and the emphasis of the old (better trained) parameters. For mixture components with high probabilistic counts, $\alpha_i^\rho \to 1$, causing the use of the new class-dependent parameters. The relevance factor is a way of controlling how much new data should be observed in a mixture before the new parameters begin replacing the old parameters. This approach should thus be robust to limited training data.

Related Entries

- ▶ Session Effects on Speaker Modeling
- ▶ Speaker Matching
- ▶ Speaker Recognition, Overview
- ▶ Universal Background Models

Acknowledgment

This work was sponsored by the Department of Defense under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

References

1. Benesty, J., Sondhi, M., Huang, Y. (eds.): Springer Handbook of Speech Processing, vol. XXXVI. Springer, Berlin (2008)
2. Müler, C. (ed.): Speaker Classification I: Fundamentals, Features, and Methods, vol. 4343/2007. Springer: Lecture Notes in Computer Science, Berlin (2007)
3. Gray, R.: Vector qantization. In: IEEE ASSP Magazine, pp. 4–29 (1984)
4. Reynolds, D.A.: A Gaussian Mixture Modeling Approach to Text-Independent Speaker Identification. PhD thesis, Georgia Institute of Technology (1992)
5. Reynolds, D.A., Rose, R.C.: Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Trans. Acoust. Speech Signal Process. **3**(1), 72–83 (1995)
6. McLachlan, G. (ed.): Mixture Models. Marcel Dekker, New York, NY (1988)
7. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. Royal Stat. Soc. **39**(1), 1–38 (1977)
8. Reynolds, D.A., Quatieri, T.F., Dunn, R.B.: Speaker verification using adapted Gaussian mixture models. Digital Signal Process. **10**(1), 19–41 (2000)

GC

GC is an analytical chemistry separation technique which provides separation of mixtures on the basis of differential affinity between a liquid or solid stationary phase and a gas mobile phase.

- ▶ Odor Biometrics

Gelatin Pad

Gelatin lifting pads are designed for the lifting of fingerprints, footprints, dust marks, and trace evidences. They comprise three layers, the first layer is the carrier, which holds the second layer of thick low-adhesive gelatin in a pliable and flexible format. The thick gelatin layer is ideal for lifting evidence without sticking to the surrounding lift area. The third layer, a cover sheet, is a clear polyester film which is removed prior to lifting, and may be replaced once the lift is completed.

- ▶ Footwear Recognition

General Model

► Universal Background Models

Generalization

The classifier is designed to correctly classify unseen objects which are not used during the training process. Generalization represents the capacity of the classifier to respond to this task. When a classifier has a good generalization capacity, it can correctly classify unseen examples.

► Ensemble Learning
► Support Vector Machine

Generalization Error

The generalization error of a machine learning model is a function that measures how far the student machine is from the teacher machine in average over the entire set of possible data that can be generated by the teacher after each iteration of the learning process. It has this name because this function indicates the capacity of a machine that learns with the specified algorithm to infer a rule (or generalize) that is used by the teacher machine to generate data based only on a few examples.

► Image Pattern Recognition

Generative Classifier

A generative classifier is a classification algorithm that learns the full joint distribution of class and attribute values. As a result, it can generate labeled instances according to this distribution. To classify an unlabeled instance, one commonly uses the Bayes decision theory.

► Fusion, Quality-Based

Genetic Identification

Identification of a victim based on the victim's DNA samples.

► Dental Biometrics

Genuine Matching

Genuine matching is matching of two templates generated from the same finger.

► Fingerprint Matching, Automatic
► Individuality of Fingerprints

Genuine Sign

Genuine sign, also called genuine signature, is a legal sign. It is legally accepted as the registered sign.

► Signature Matching

Genuine/Impostor Attempt

In a genuine attempt, a biometric sample is compared against other biometric samples from the same subject. If similarity between the samples is not high enough, the subject will be wrongly rejected by the system. In an impostor attempt, a biometric sample is compared against biometric samples from other subjects.

If similarity between the samples is high enough, the subject will be wrongly accepted by the system. It should be noted that biometric samples from the same user are not necessarily similar (e.g., temporary injuries in the finger) and on the other hand, biometric samples from different users can be quite similar (e.g., signature forgeries).

► Fingerprint Databases and Evaluation

Geodesic

Geodesic is the integral curve between two points corresponding to the gradient direction of the intrinsic distance function of the manifold.

► Manifold Learning

Geometry Image

A geometry image is the result of representing all vertices of a 3D object (x, y, and z coordinates) as a simple 2D array of quantized points. Geometry images have at least three channels assigned to each u, v pair of coordinates, encoding geometric information (x, y, z coordinates) of a vertex in R^3, but surface normals and colors can also be stored using the same implicit surface parametrization. Creating a geometry image is accomplished by cutting an arbitrary mesh along a network of edge paths and parametrizing the resulting single chart onto a square.

► Face Recognition, 3D-Based

Global Fusion

Global fusion in the framework or multi-biometric score fusion refers to user-independent score fusion techniques in which a unique fusion function is used for all users, which is trained based on background data from a pool of users (both genuine and impostor scores).

► Fusion, User-Specific

Global Thresholding Techniques

Global thresholding technique is used to convert an image consisting of gray scale pixels to one containing only black and white pixels. Usually a pixel value of 0 represents white and the value 255 represents black with the numbers from 1 to 254 representing different grey levels. A threshold value Th is chosen in the range of 1–254 and each grey pixel P in the image is modified to either black or white according to the test.

If $P \geq Th$ then $P = 255$ (white) orelse $P = 0$ (black).

There are a number of ways to select the value of threshold Th depending on the nature of grey pixel distributions in the image.

► Hand Vein

Glottal Excitation

The glottal excitation corresponds to the pulsating flow of air that comes from the lungs through the vibrating vocal folds. This first process of the human speech production mechanism is named after the orifice between the vocal folds, the glottis.

► Speech Production

Glyph

A glyph is the shape of a handwriting sample. In Roman scripts, it may contain one letter or even a

group of letters depending on the content of the sample. In oriental scripts, a glyph corresponds to a character which consists of a set of strokes.

► Signature Sample Synthesis

GMM

► Gaussian Mixture Models

Graph Matching

The configural identification of a face relating to the measurable distances between features and the relative ratios of height and width. A unique algorithm is created from the key points on the face; this algorithm is regarded as a unique biometric identifier.

► Face, Forensic Evidence of

Graphic Tablet

► Digitizing Tablet

Graphical User Interface

► User Interface, System Design

Graphometric Features

Graphometric features are intrinsic properties from an individual handwriting style, which may be employed by forensic experts during handwriting or signature recognition. These include curvature and pressure among others.

► Signature Features

Gray Scale

A continuous-tone image that has one component, which is luminescent.

► Vascular Image Data Format, Standardization

GRF (Ground Reaction Force)

The ground reaction force is, according to Newton's law of reaction, the force equal in magnitude but opposite in direction produced from the ground as the reaction to force the body exerts on the ground. The ground reaction force is used as propulsion to initiate and control the movement, and is normally measured by force sensor plates.

► Footstep Recognition

Ground-Truth

The actual facts of a situation, without errors introduced by sensors, software processing or human perception and judgment. For example the actual location of a minutia in a fingerprint image that could be used to check the accuracy of the location reported by a given automated minutiae extraction algorithm.

► SFinGe

Gummy Bear Finger

► Fingerprint Fake Detection

H

Haar-Like Features

Similar to the what Haar wavelets are developed for basis functions to encode signals, the objective of two-dimensional Haar features is to collect local oriented intensity difference at different scale for representing image patters. This representation transforms an image from pixel space to the space of wavelet coefficients with an over-complete dictionary of features. Such features can be used to represent face and pedestrians images. The Haar-like features, similar to Haar wavelets, compute local oriented intensity difference using rectangular blocks (rather than pixels) which can be computed efficiently with an integral image.

► Face Detection

Habituated Subject

A user of a biometric system who is well versed in its use; someone who routinely uses a biometric system.

► Iris on the Move

Habituation

The academic and medical world has several different definitions for habituation. Two recurrent characteristics in the literature are acclimation and habituation. The first is acclimation, which consists of a user's first exposure to a device or process, the formal training that goes along with it, and individual learning and experimentation. The second characteristic, habituation consists of two parts, partially habituated and fully habituated.

1. Acclimation is the process in which a user of a biometric system adapts his or her techniques to achieve a proper match of his or her biometric template.
2. External Teaching is the formal training that a user receives revealing proper techniques and the series of steps included with using the biometric system.
3. Self teaching occurs after external teaching where a user experiments with the device and begins to eliminate techniques that do not work well or are not comfortable. Through this iterative process, techniques that work are narrowed, leading to partial habituation.
4. Partial habituation is defined as the period of time when no new adaptation of techniques is used to achieve a proper match to the biometric template.
5. Full or complete habituation is defined as the point at which a user matches his or her biometric template using subconscious techniques.

► Ergonomic Design for Biometric Systems

Halo Effect

The temperature difference between a wet finger and the platen of an optical sensor generates a halo on the final image around the fingerprint.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

© 2009 Springer Science+Business Media, LLC

Hamming Distance

A measurement of the (dis)similarity between two strings of bits having equal length, based on tallying how many corresponding pairs of bits in the two strings disagree. If each string has N bits, then a *cardinal* Hamming Distance is the count of disagreeing bits and is thus an integer between 0 and N inclusively. Alternatively, a *fractional* Hamming Distance normalizes (divides) this count by the total N and is thus a rational number between 0 and 1. Hamming Distance is an extremely fast metric to compute because it can be implemented digitally by simple Exclusive-OR logic operating in parallel on chunks of bits as large as the word length of the processor itself (*e.g.*, 64 bits at once) in a single executable instruction cycle, and thus within almost a single "tick" of the system clock. In dedicated hardware there is no necessary limit to how many bits can be XOR'ed at once, thus allowing Hamming Distances to be computed at virtually unlimited rates. This confers an advantage to this similarity metric when searching databases on a national scale. A normalized Hamming Distance is the metric underlying the matching of IrisCodes for recognizing persons by their iris patterns.

► Iris Encoding and Recognition using Gabor Wavelets
► Score Normalization Rules in Iris Recognition

Hand Biometrics

► Hand Geometry

Hand Biometrics, 3D

► Finger Geometry, 3D

Hand Contour

► Hand Shape

Hand Data Interchange Format, Standardization

Raul Sanchez-Reillo[1], Samir Tamer[2]
[1]University Carlos III of Madrid, Avda. Universidad, Leganes (Madrid), Spain
[2]Ingersoll Rand Recognition Systems, Dell Avenue, Campbell, CA, USA

Synonyms

Encoding of hand geometry information; Hand silhouette data

Definition

Standard that defines a common format to code information related to hand geometry based biometrics. This format is defined to allow interoperability among different vendors worldwide, and has been developed by the international community taking part in ISO/IEC JTC1/SC37 standardization subcommittee.

Introduction

Subcommittee SC37 from ISO/IEC JTC1 deals with the standardization of biometrics. Among SC37 Working Group 3 is devoted to define Interchange Data Formats for biometric modalities, among other duties. For that purpose, a multipart standard is under development, and it is referred to by the number ISO/IEC 19794. Part 10 of the multipart standard covers hand geometry biometrics, and is denoted ISO/IEC 19794-10. The full title is "Information technology - Biometric data interchange formats - Part 10: Hand geometry silhouette data"[1].

This International Standard provides a data interchange format, based on a CBEFF data block [2], for applications requiring an interoperable hand geometry record. The information consists of a variety of mandatory and optional items, including data capture parameters, standardized hand position, and vendor specific information. This information is intended for interchange among organizations that rely on automated devices and systems for identification or verification purposes based on the information from hand geometry.

It is important to note that although this part of ISO/IEC 19794 mandates a particular data format, it does not mandate a particular algorithm. For example, a user may be enrolled on a system from one vendor, and verified on a system from another.

Also, an important issue is that this format stores hand silhouette data rather than color or greyscale image data. To increase the flexibility of the data format, provisions have been made to store views of the left and right hands, in addition to multiple views of each hand.

Specific implementations of this part of ISO/IEC 19794 that could be constrained by storage space or transmission capability (such as smart card applications) may wish to limit the number of views stored for each hand. Such limitations are outside the scope of this part of ISO/IEC 19794, but authors of the International Standard advise that the reduced choices can prejudice interoperability.

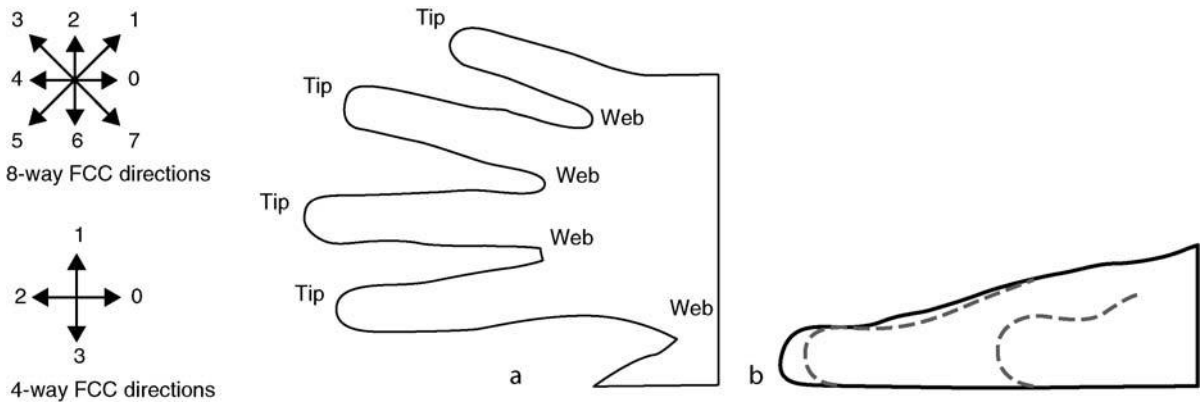
Silhouette Acquisition Requirements

The capture device as well as the capture process is out of the scope of the standard. As already mentioned, this is not an image-based standard, but one related to the coding of the shape of the hand. Therefore, no matter what camera has been used for acquiring the sample (black and white, colour, resolutions, etc.), or which algorithm has been used for preprocessing such image, the primary input for this document is such preprocessed image, showing the silhouette of the hand captured. This silhouette can be either the one referring to the top-view of the hand, or its side-view. Figure 1 shows the standardized orientation of both types of view.

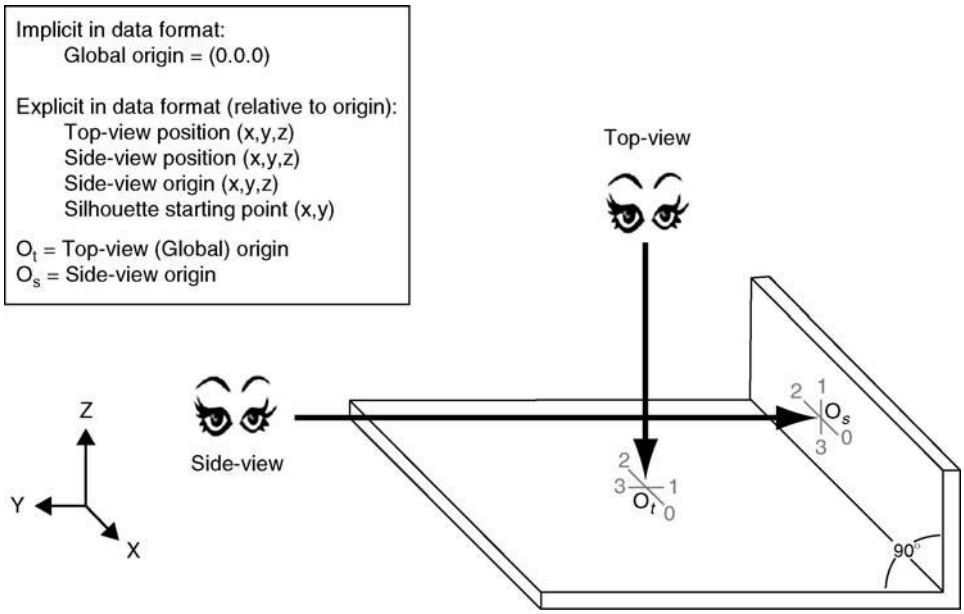
The hand silhouette will be represented in this standard as a sequence of points showing the direction to the next point in the silhouette (what is called a ► Freeman Chain Code or ► FCC). In order to code the FCC in an interoperable way, a set of requirements have to be defined:

- The basic requirement is that aspect ratio shall be 1:1, with an error less than ±2%
- The starting point shall be in the rightmost column of the silhouette in Fig. 1, at the uppermost row occupied by the silhouette in that column (i.e., the upper right corner of the silhouette). Successive points shall trace the outline in a counter-clockwise direction.
- The silhouette shall be a closed shape (i.e., it shall have no gaps in the outline, and the final outline point shall be common to the starting point).
- The starting point shall occur exactly twice in the silhouette, as the first and last points only (the silhouette will not cross through the starting point at any other time)
- The right column shall be vertical (i.e., the penultimate point shall occur directly below the starting point, and no points occur to the right of the starting point)

The orientation of the camera while capturing the image is quite important. Figure 2, shows the coordinate system when dealing with the



Hand Data Interchange Format, Standardization. Figure 1 Standard hand orientation images: (a) top-view, (b) side-view. Images taken from [1].



Hand Data Interchange Format, Standardization. **Figure 2** Coordinate System linking top and side views, referred to the 4-FCC directions. Image taken from [1].

► camera point of view. This is really significant for some data to be stated in the record to be coded.

Record Format

After defining the set of requirements for image acquisition, the Standard defines the way such information has to be coded and stored within a CBEFF-compliant wrapper. The structure to be followed is:

- A fixed-length (15-byte) general record header containing information aboutthe overall record, with the following fields:
  - Format identifier (4 bytes with the hexadecimal value 0x484E4400) and version number (coded in another 4 bytes)
  - Record length (in bytes) including all hand views within this record (coded in 4 bytes)
  - Number of hand views (HGVRs) (1 byte)
  - 2 bytes reserved for future use
- One or more variable-length ► Hand Geometry View Records (HGVRs), eachcontaining a single hand silhouette, consisting of:
  - A fixed-length (25-byte) hand view header contains the following information:
    - Length of the HGVR (in 2 bytes)
    - HGVR index (in 1 byte)
    - Hand identifier (1 byte), which indicates the fingers that the system attempts to acquire within the silhouette, and the view of the hand (top view of the palm, top view of the back of the hand, side view from the thumb side, or side view from the little finger side)
    - Hand Integrity (1 byte), which shows the identified problems in the sample acquired (e.g., finger missing, misplacement, etc.)
    - Data resolution in pixels per centimetre (1 byte)
    - Geometric distortion of the system, as a signed value incrementing 0.1% (1 byte)
    - Silhouette quality (3 bytes), with being 0 the lowest quality and 100 the highest possible quality, always coded in the lower byte, while the higher 2 bytes are reserved for the future use
    - Camera position relative to the global origin (1 byte for X position, 1 byte for Y position, and 1 byte for Z position)
    - Target position relative to the global origin (1 byte for X position, 1 byte for Y position, and 1 byte for Z position)
    - Silhouette starting point relative to the view origin (1 byte for X position and 1 byte for Y position)
    - Data compression algorithm (1 byte), which currently refers to only 2 coding methods, such as 8-way FCC and 4-way FCC

    - Hand scanning technology (1 byte), giving information whether the image was acquired using an optical camera, a linear scanning array, or no information is specified
    - Extended data length (2 bytes)
    - 3 bytes reserved for future use
  - Silhouette data, encoded using a Freeman Chain Code (FCC), either using 8-way FCC, or 4-way FCC (depending on what is declared in the "Data compression algorithm" field at the HGVR header
  - Extended data (optional), for any application-specific or proprietary data used by the system vendor

For further details refer to the current version of this International Standard [1]. Current version also provides a record sample, as well as an informative annex related to the best practices in this biometric modality, including hand placement and platen and optical design.

Other Related Standards

There are other standards related to this technology, born under ANSI/INCITS scope. This is ANSI/INCITS 396-2005 – "Information Technology - Hand Geometry Interchange Format." This standard is extremely similar to ISO/IEC 19794-10, where the major technical differences are:

- Within the General Header:
  - ANSI/INCITS 396 includes a CBEFF Product Identifier
  - ANSI/INCITS 396 version number is a binary byte, while in 19794-10 is a 4-byte string
- Regarding the View Header:
  - ANSI/INCITS 396 has a creation date that was dropped by ISO 19794-10
  - 19794-10 adds a view index that associates multiple views of the same hand (such as a top-view and side-view captured at the same time)
  - 19794-10 adds a Hand Integrity field that indicates which fingers are ok and which are missing/mangled
  - 19794-10 adds a starting-point location linking the absolute position of the silhouette to the camera's optical axis
  - 19794-10 supports 4-way or 8-way FCCs, where ANSI/INCITS 396 only supports 8-way

Due to the fact that the International Standard ISO/IEC 19794-10 is already available, INCITS has withdrawn ANSI/INCITS 396-2005.

Summary

To provide interoperability in storing and transmitting hand-geometry-related biometric information, one international standard has been developed. Beyond this International Standard, other standards deal with conformance and quality control, as well as interfaces or performance evaluation and reporting (see relevant entries in this Encyclopedia for further information).

Related Entries

- ▶ Biometric Data Interchange Format
- ▶ Common Biometric Exchange Framework Formats
- ▶ Hand Geometry
- ▶ Hand-Geometry Device
- ▶ Palm Vein

References

1. ISO/IEC: 19794-10:2007 - information technology - biometric data interchange formats - part 10: Hand geometry silhouette data (2007)
2. ISO/IEC: 19785-1:2005 - information technology - common biometric exchange formats framework - part 1: Data element specification (2005)

Hand Databases and Evaluation

Guangming Lu
School of Computer Science and Technology, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China

Synonyms

Hand geometry; Hand vein; Palmprint database; Palm vein

Introduction

The human hand is the source of a number of unique physiological characteristics. The main technologies for hand recognition fall into three categories: palmprint technologies – those measuring the unique pattern of prints from the palm of the hand – similar to a fingerprint; Hand geometry measurements – those measuring the shape and size of either all or part of the human hand or fingers; Hand vein patterns – those measuring the distinct vascular patterns of the human hand, including hand dorsum vein and palm vein.

Palmprint: A palmprint is defined as the skin patterns of a palm, composed of the physical characteristics of the skin patterns such as lines, points, and texture. Palmprints are covered with the same type of skin as a finger, but their surface area is much larger than a finger tip [1]. Automatic palmprint identification systems can be classified into two categories: online and offline. An online system captures palmprint images using a palmprint capture sensor that is directly connected to a computer for real-time processing. An offline palmprint identification system usually processes previously captured palmprint images, which are often obtained from inked palmprints that have been digitized by a digital scanner. Several offline palmprint databases were set up for algorithms design during the early stages of palmprint research and they yielded promising results. Today most researchers focus on the online palmprint recognition techniques and system development. Some public online palmprint databases are readily available for the researcher to design efficient recognition algorithms.

Hand geometry: The hand geometry biometric involves several features including the length, width, thickness, and surface area of the hand or fingers of a user [2]. Hand geometry has several advantages over other biometrics, including small feature size, and smaller computation requirement as a result of using low resolution images. In spite of its current widespread use (4.7% market share in 2007), the hand geometry system suffers from high cost and low accuracy. In addition, uniqueness of the hand features is not guaranteed, making it unsuitable for use in one-to-many identification applications [1].

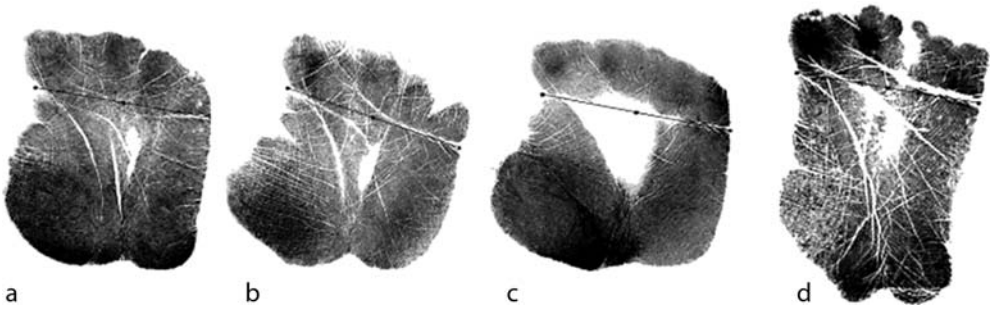
Hand dorsum vein: The idea of verifying user identity on the basis of the pattern of subcutaneous veins on the back of the hand was first proposed by MacGregor and Welford in 1991 [3]. A palm dorsum vein pattern authentication system named VP-II is provided by the company I-OnAsia. The system uses an infrared sensor to capture the thermal image of the hand dorsum vein pattern. Since no contact with the device is required, it ensures excellent convenience, sanitary use, and prevents copying of system-residual biometrics.

Palm vein patterns: The Fujitsu Laboratories Limited developed a new type of biometric authentication technology which verifies a person's identity on the basis of the pattern of veins in his/her palms [4]. It makes use of the characteristic of light absorption difference for oxygenated and deoxidized hemoglobin in the blood to capture the pattern of the palm vein. The image of the palm vein would appear as a black pattern if the palm is illuminated with a near-infrared light. This is because some of the infrared light is absorbed by the vein and thus weakens the reflected light back. The company claims that, except for their size, palm vein patterns are personally unique, and do not vary over the course of a person's lifetime.

Palmprint Databases and Evaluation

The earliest paper on palmprint identification [5] reported in detail the construction of an offline palmprint database, which included palmprints of 20 individuals – 10 prints from both left and right palms yielding a total of 400 palmprint images. The authors used 100 dpi resolution to digitize these inked palmprints paper sheets with 400×400 pixels to reduce the computation burden. A sub database, which included 60 special palmprint images, was employed to ascertain the performance of palmprint line matching algorithm using datum points. This sub database consisted of 20 rotated palmprint images, 35 partial palmprint images, and 5 palmprint images in which the life line and head line did not touch each other. Some typical palmprint images are shown in Fig. 1 and the results of the datum point detection from special palmprint images are given in Table 1.

Another offline palmprint database was developed and reported by You, Li, and Zhang [6]. It was collected from 100 individuals with two prints from the right palm, and comprised a total of 200 palmprint images. The palmprint images were of 232×232 pixels, with the resolution of 125 dpi and 256 gray levels. The palmprint samples were collected from both female



Hand Databases and Evaluation. Figure 1 Examples of datum points determination: (**a**) normal palmprint, (**b**) rotated palmprint, (**c**) incomplete palmprint, and (**d**) the life line and head line unintersection [5].

Hand Databases and Evaluation. Table 1 Experimental results of datum point determination in special palmprint images

Classification	R	I	U
Experiment images	20	35	5
Accurate determination images	19	33	4
Rate of accuracy (%)	95	94	80

R rotated image; *I* incomplete image; *U* life line and head line unintersection

and male adults in the age group of 18–56 years. Samples of such palmprint images are shown in Fig. 2. They proposed hierarchical palmprint identification via multiple feature extraction. First, a texture based dynamic selection scheme is proposed to facilitate rapid search for the best matching of the sample in the database. The global texture energy is used to guide the dynamic selection of a small set of similar candidates from the database at a coarse level for further processing. An interesting point based image matching is performed on the selected similar patterns at a fine level for the final confirmation. The experimental results from this database show that the average accuracy rate is 95%. Since the majority of samples was filtered out by the coarse level classification, the execution speed of the fine level feature matching increased significantly.

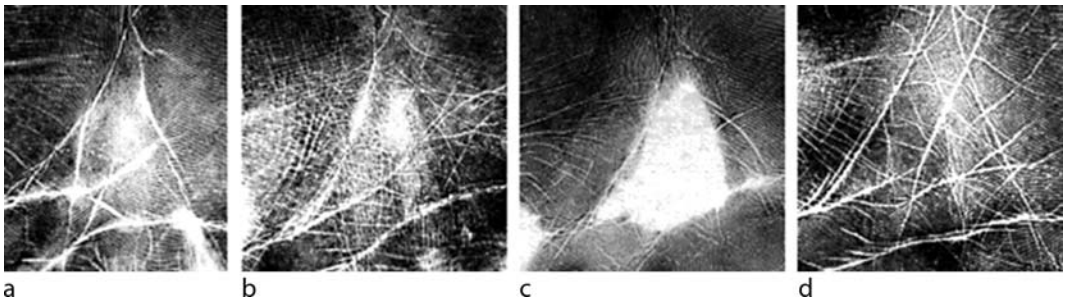
At present many different online palmprint databases have been used in research papers. Among these [7, 8, 9], the following two public databases are more representative: The Hong Kong Polytechnic University (PolyU) Palmprint Database 1.0 and 2.0 [9]. The PolyU Palmprint Database 1.0 is a subset of 2.0.

The PolyU Palmprint Database 1.0 contains 600 palmprints which were collected from 100 different palms. Six samples from each of these palms were collected in two sessions, in other words, three samples were captured in each session. The palmprints were of 384 × 284 pixels and with 256 grayscales. Some-typical palmprint images in this database are shown in Fig. 3. Many research papers have been published that frequently employ this database. The performance achieved in the literature is very high and comparable with any other established biometric modality. For example, Wangmeng Zuo et al. [10] achieved a recognition rate of 97.67% by using 2DPCA-AMD.

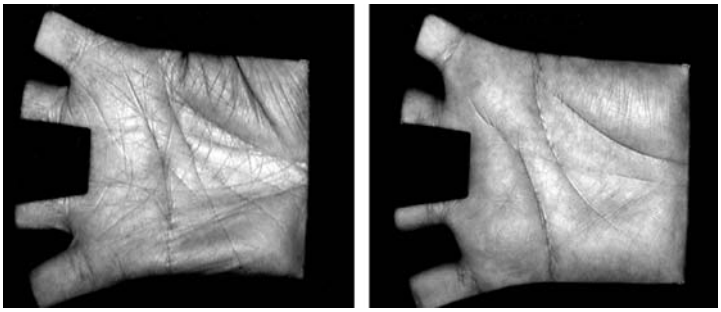
The PolyU Palmprint Database 2.0 contains 7,752 grayscale images corresponding to 386 different palms, and all the images are available in bitmap format. Around 20 samples from each of these palms were collected in two sessions, that is around 10 samples were captured in the first session and the second session, respectively. The average interval between the first and the second collection was 2 months. In this dataset, there were 131 males, and the age distribution of the subjects was as follows: younger than 30 years about 86%, older than 50 about 3%, and around 11% between 30 and 50 years. Numerous papers in the literature have employed the PolyU Palmprint Database 2.0. The typical algorithms that revealed most promising results are Competitive Coding [11], RLOC (Robust Line Orientation Coding) [12], and DoG (Derivative of Gaussian Coding) [13]. The testing results of the three methods based on the PolyU Palmprint Database 2.0 are given in Fig. 4.

Hand Geometry Databases and Evaluation

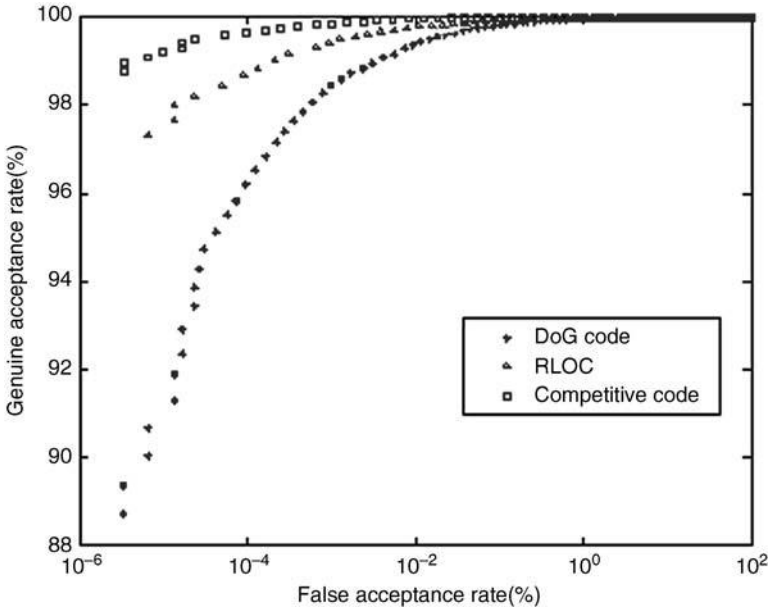
The hand geometry features can be typically extracted from a binary image. Therefore, low-resolution



Hand Databases and Evaluation. **Figure 2** Samples of different palmprint patterns with distinctive texture features (**a**) strong principle lines, (**b**) full of wrinkles, (**c**) less wrinkles, and (**d**) strong wrinkles [6].



Hand Databases and Evaluation. **Figure 3** Some typical palmprint images from PolyU palmprint database 1.0.



Hand Databases and Evaluation. **Figure 4** The ROC curves of the three typical identification methods based on the PolyU palmprint database 2.0.

imaging is employed to gain the advantage associated with faster processing. The hand geometry database introduced by Savic and Pavesi [14], includes the gray-scale hand images of 100 individuals (68 males and 32 females) with 10 images of the right hand and 10 images of the left hand (a total of 2,000 images). The average age of the subjects was 36 years; the oldest was 78 years and the youngest was 21 years, all subjects belonged to the same ethnic group. The images were collected in two separate sessions over a period of 3 months. As the left and right hands are different, the left hand images were mirrored and used as the right hand images of "new" people. In this way, 200 image classes with 10 images per class were obtained. The geometrical parameters extracted from the hand are illustrated in Fig. 5(a), and the experimental result is presented in Fig. 5(b).

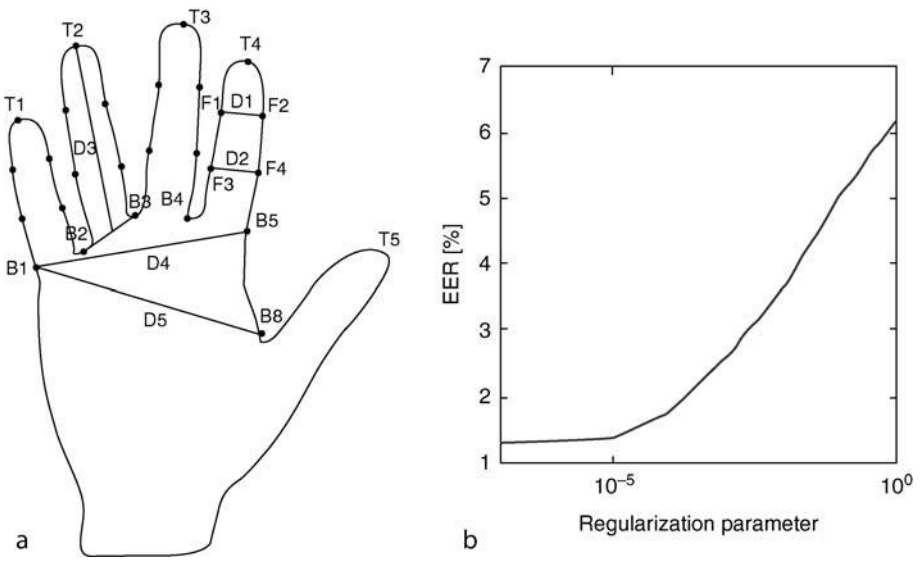
Kumar et al. [15] developed a hand database captured by a digital camera, in which they collected 1,000 images of the left hand from 100 subjects, i.e., 10 samples from each subject. The geometry features are defined in Fig. 6(a). Based on this method, the testing results are presented in Fig. 6(b) when they used the first five images from each subject for training and the remaining for testing.

One of the earliest hand geometry database construction was reported by Sanches-Reillo et al. [16], which comprises 200 hand images obtained from 20 people of different ages, sex, profession, and living style. Ten samples were collected from each person on different days over a period of 3 months. Experimental results show that the proposed method can achieve up to 97% accuracy in classification by using Gaussian Mixture Modeling.

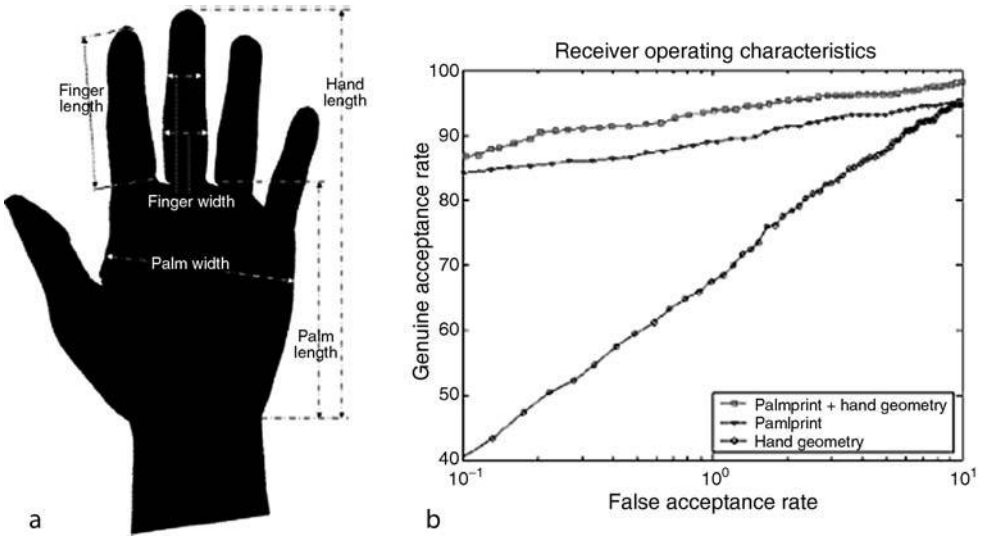
Hand Dorsum Vein Databases and Evaluation

Hand dorsum vein pattern has been widely researched during the last decade. Some typical hand dorsum vein images obtained from low-cost near infrared imaging are shown in Fig. 7. The researchers designed their own capture devices to collect hand dorsum vein images. So far, there is no public hand dorsum vein database available to researchers. Here, the attempt is to introduce some typical hand dorsum databases from published research papers.

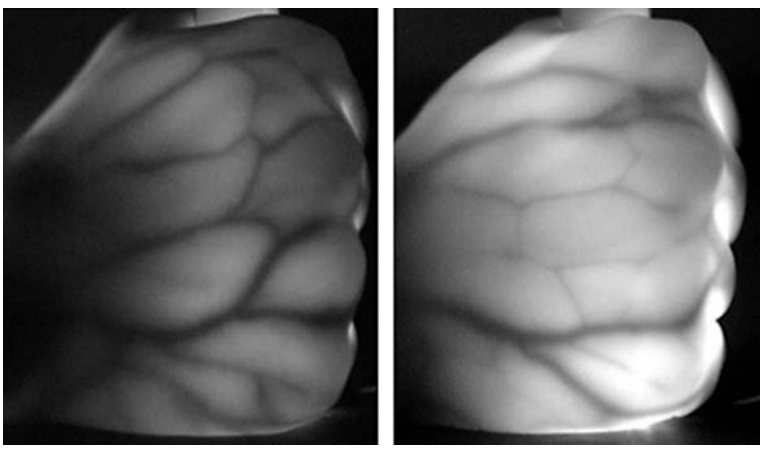
Lin and Fan [17] obtained 960 palm dorsa thermal images from 32 volunteers. The volunteers in this set of verification experiments included 29 male adults and 3 females. Each of the thermal images of 640 × 480 pixels was captured under varying light conditions (even in the dark). The authors obtained an equal error rate of 2.3%. Another hand dorsum vein database which includes 300 images from 100 persons



Hand Databases and Evaluation. Figure 5 (**a**) Geometrical parameters of the hand, (**b**) the EER as a function of the regularization parameter [14].



Hand Databases and Evaluation. Figure 6 (**a**) Defined geometrical features, (**b**) the ROC of the testing results (note the hand geometry curve) [15].



Hand Databases and Evaluation. Figure 7 Typical hand dorsum vein images.

of different ages and gender has been employed by Kumar and Prathyusha [18]. This database was developed by using a low-cost near infrared camera, traditionally utilized for surveillance, using contactless image acquisition. The volunteers were requested to present their folded right hand (palm dorsal surface) near the imaging window such that the knuckle tip of the middle finger remained at the top. The authors [18] obtained an equal error rate of 1.14% by integrating knuckle shape information into this contactless image database. Sang-Kyun Im et al. [19] developed a hand dorsum vein database comprising 5,000 grayscale images. Each image was of 160 × 120 pixels. They evaluated the database by using CSD codes and achieved a reliability accuracy of 99.45%. By using the FPGA (field programmable gate array) device, the whole verification process required only 150 ms per person.

Palm Vein Database and Evaluation

Research on palm vein technologies is fairly recent and employs low-cost near infrared imaging to obtain palm vein images. The Fujitsu Laboratories Limited developed a palm vein system by testing 1,400 palm profiles of

700 individuals, the system achieved a false rejection rate of 1% and a false acceptance rate of 0.5% with an equal error rate of 0.8% [4]. The palm vein image database used by Wang et al. [20] was constructed from 120 subjects and each image was of 768 × 576 pixels. This database also contains palmprint images as the employed image acquisition device can simultaneously acquire palm vein images from near infrared imaging and palmprint images from visible illumination.

Summary

Hand based biometric technologies have attracted a lot of attention both in research and industry. An increasing number of hand based biometric systems have appeared in the literature. There is a lot of interest in research on hand based biometrics for the simultaneous acquisition of several hand based traits. These multimodal biometric technologies try to combine all the features (palmprint, hand geometry, hand vein, palm vein) of the hand to build a higher accuracy biometric system. Such multimodal systems offer very high resistance against spoof attacks and have diverse applications.

Related Entries

- ▶ Hand Geometry
- ▶ Hand Vein
- ▶ Palm Vein
- ▶ Palmprint

References

1. Zhang, D.: Palmprint Authentication, Kluwer, Dordrecht (2004)
2. Kumar, A., Zhang, D.: Hand geometry recognition using entropy-based discretization. IEEE Trans. Inf. Forensics and Secur. **2**(2), 181–187 (2007)
3. MacGregor, P., Welford, R., Veincheck, R.: Imaging for security and personnel identification. Adv. Imaging. **6**(7), 52–56 (1991)
4. Fujitsu Laboratories Limited. Contactless Palm Vein Pattern Biometric Authentication System. http://pr.fujitsu.com/en/news/2003/03/31.html
5. Zhang, D., Shu, W.: Two novel characteristics in palmprint verification: datum point invariance and line feature matching. Pattern Recogn. **32**(4), 691–702 (1999)
6. You, J., Li, W., Zhang, D.: Hierarchical palmprint identification via multiple feature extraction. Pattern Recogn. **35**(4), 847–859 (2002)
7. http://visgraph.cs.ust.hk/biometrics/Visgraph_web/index.html. Last Accessed 27 March, 2009
8. http://www.cbsr.ia.ac.cn/china/Palmprint%20Databases%20CH.asp. Last Accessed 27 March, 2009
9. http://www.comp.polyu.edu.hk/~biometrics. Last Accessed 27 March, 2009
10. Zuo, W.M., Zhang, D., Wang, K.Q.: Bidirectional PCA with assembled matrix distance metric for image recognition. IEEE T. Syst. Man Cy. B. **36**(4), 863–871 (2006)
11. Kong, A.W.K., Zhang, D.: Competitive coding scheme for palmprint verification. Pattern Recogn. **1**, 520–523 (2004)
12. Jia, W., Huang, D., Zhang, D.: Palmprint verification based on robust line orientation code. Pattern Recogn. **41**(5), 1054–1513 (2008)
13. Wu, X., Wang, K., Zhang, D.: Palmprint texture analysis using derivative of gaussian filters. In: 2006 International Conference on Computational Intelligence and Security, Berlin, 751–754 (2006)
14. Savic, T., Pavesi, N.: Personal recognition based on an image of the palmar surface of the hand. Pattern Recogn. **40**, 3152–3163 (2007)
15. Kumar, A., Wong, D.C.M., Shen, H.C., Jain, A.K.: Personal authentication using hand images. Pattern Recogn. Lett. **27**, 1478–1486 (2006)
16. Sanches-Reillo, R., Sanchez-Avila, C., Gonzalez-Marcos, A.: Biometric identification through hand geometry measurements. IEEE T. Pattern Anal. **22**(10), 1168–1171 (2000)
17. Chih-Lung, L., Kuo-Chin, F.: Biometric verification using thermal images of palm-dorsa vein patterns. IEEE T. Circ. Syst. Vid. **14**(2), 199–213 (2004)
18. Kumar, A., Prathyusha, V.: Personal authentication using hand vein triangulation. In Proceedings of the SPIE, vol. 6944, pp. 69440E–69440E–13 (2008)
19. Sang-Kyun, I., Hyung-Man, P., Young-Woo, K., Sang-Chan, H., Soo-Won, K., Chul-Hee, K.: A biometric identification system by extracting hand vein patterns. J. Korean Phys. Soc. **38**(3), 268–272 (2001)
20. Wang, J.-G., Yau, W.-Y., Suwandy, A., Sung, E.: Person recognition by fusing palmprint and palm vein images based on laplacianpalm representation. Pattern Recogn. **41**(5), 1514–1527 (2008)

Hand Geometry

Raul Sanchez-Reillo
University Carlos III of Madrid, Avda. Universidad, Madrid, Spain

Synonyms

Hand biometrics; Hand shape biometrics

Definition

Biometric modality based on identifying a person by the shape of his or her hand. In its basic form, it is based on taking a photograph of the user's hand while placed on a surface, and after a ► contour detection, finding singular points and taking measurements among them.

Introduction

Hand Geometry is considered as a medium-profile biometric modality which reaches a really high level of user acceptance with low computational cost. Not being one of the first biometric modalities, it has gained great popularity due to the success of some commercial products, at the end of the twentieth century. In fact, the commercial product from Schlage Recognition Systems, known as HandKey II [1], was one of the most sold at the beginning of the 2000s, especially for physical access control systems and time and attendance control.

As mentioned below, after some initial works, other scientists have continued researching on other algorithms and more comfortable means of using this technology. Error rates achieved are not as low as those modalities considered as high-performance ones (e.g., fingerprint, iris or vascular). In order to gain applicability, some researchers have included this technology in multimodal biometric systems, reducing error rates, and gaining in ► usability and user acceptance.

Basics and Initial Works

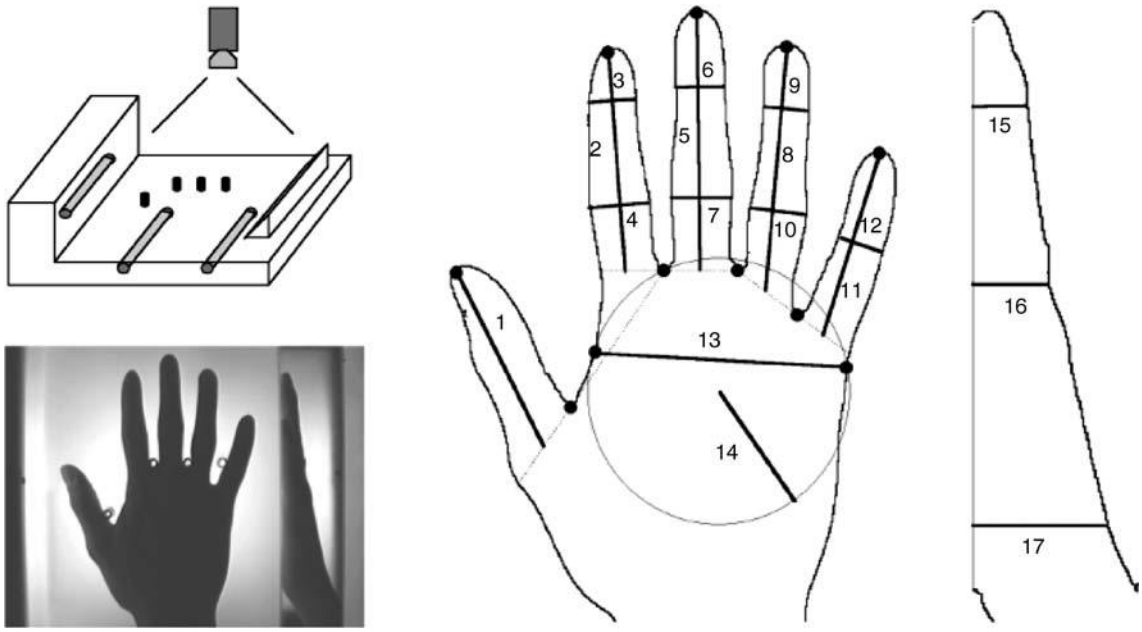
Hand Geometry biometrics is based on the measurement of the shape of the contour of the hand [2], including finger widths at several points, finger lengths, palm shape, deviation angles, etc. Main idea comes from the Bertillon system (http://en.wikipedia.org/wiki/Bertillon) used during the late nineteenth century to identify prisoners. But it was not till 1997 that the first paper in a scientific journal is found. In such paper, among may other interesting things, Golfarelli et al. [3] outline a system based on a semi-opaque plastic material with some fixed ► pegs to guide positioning of the hand. With a CCD camera located over the hand, and some light located under the surface plate, a high contrast image of the user's hand is obtained. As to acquire also the lateral projection of the hand, the system is replicated on the side, and a 45 mirror is placed to project such image to the same camera. The counterlight image allows a very easy contour detection of the hand, and from there 17 geometrical features are extracted. Figure 2 illustrates image acquisition and feature extraction steps from this work.

In 1999 and in 2000 two papers were published detailing this biometric modality. They were written

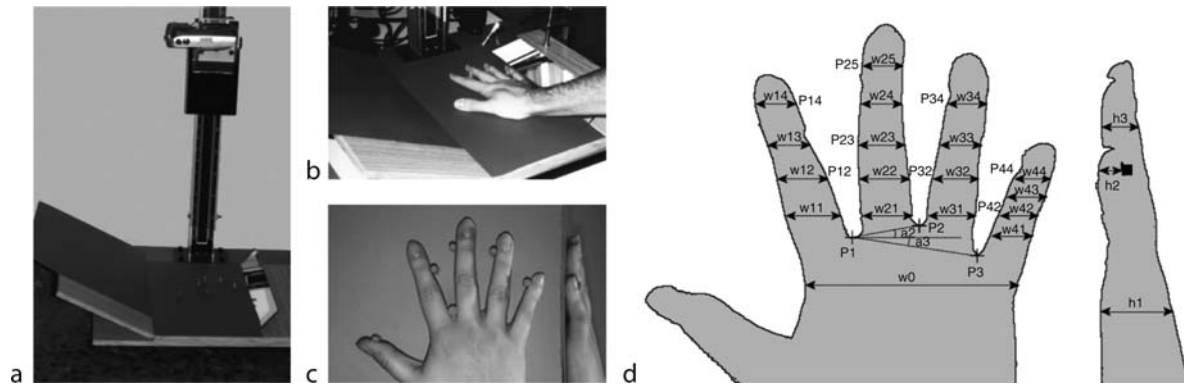
Hand Geometry. Figure 1 HandKey II device and illustration of its use as a door lock in a physical access control system [1] (Images published under authorization of Schlage Recognition Systems).

by Jain et al. [4] and Sanchez-Reillo et al. [5]. In this last work, the device developed is also based on a CCD camera located over the hand of the user. But differently from the Golfarelli approach [3], here the hand is located over an opaque peg-oriented surface painted in blue (see Fig. 3). The reason for the platform to be blue is that the human skin, no matter the race, has a very low portion of blue component. Therefore discarding all blue component in the RGB, allows an easy way to eliminate all background information. As in Golfarelli et al. system [3], also a mirror is placed to obtain the lateral view of the hand. In contrast, the illumination demands of this new system were lower, as only the one coming from the camera built-in flash was employed.

From the image sample acquired, as mentioned above, the background surface is removed by eliminating the blue component of the image. In fact, in order to obtain a better output for next stage, the background removal is done by the following formulae (cropping all negative values to 0):



Hand Geometry. **Figure 2** Illustration of the Hand Recognition system designed by Golfarelly et al., including a sample of the photographs taken and the geometrical measurements extracted. Images taken from [3]. ©IEEE.



Hand Geometry. **Figure 3** Prototype developed in [5] and measurements taken: (**a**) General view of the prototype; (**b**) Positioning of the hand; (**c**) Sample taken; (**d**) Geometric features. Images extracted from [5]. ©IEEE.

$$I_{BW} = ((I_R + I_G) - I_B) \quad (1)$$

Afterwards, a Sobel edge-detection is performed, obtaining a binary image where only the borders are in black, while the rest of the image is in white. With this result, the way to obtain the 31 absolute features is by locating singular points in the image, and counting pixels among them. In [5], authors proposed the following basic features (as seen in 3d):

- Palm width at a certain point (avoiding conflict with pegs). It is obtained as the number of pixels from the first black pixel on the right side of the top view of the hand, to the next black pixel along the same horizontal line, after going through a set of white pixels moving to the left.
- Finger widths at certain points, also avoiding pegs and ring area, and obtained in an analogue way as mentioned with the palm width.
- Palm and finger heights, through the same mechanism as mentioned above, but this time with the side view of the hand.
- Finger curvature or Deviations. This is defined by the distance between a middle point of the finger and the middle point of the straight line between the inter-finger point and the last height where the finger width is measured. Equation used can be seen below, where exponents refer to the coordinate used, and subindex defines the finger used and the width measurement used.
- Angles between inter-finger points and the horizontal line, which reflects the depth of each of the inter-finger point.

$$deviation = P_{12}^X - \frac{P_{14}^X - P_1^X}{P_{14}^Y - P_1^Y}(P_{12}^Y - P_1^Y) \quad (2)$$

From those features, the feature space was grown by adding relative measurements, i.e., relationships among different sets of basic features. Authors, after applying a Principal Component Analysis, discovered that from all those measurements, only 25 features had significant discriminant properties. Figure 3d shows the 25 absolute measurements from whose the final 25 features extracted.

Authors researched the behavior of four different comparators: Euclidean Distance, Hamming Distance in the continuous domain (as seen in equation (3) where x_i refers to the ith component of the sample, L is the feature vector length, t_i^m is the mean of the ith component, and t_i^v is the standard of the ith component), Gaussian Mixture Models (GMM) and Radial Basis Function Neural Networks (RBF-NN). Also they analyzed the dependence of the performance with the number of samples used during enrollment. Results showed that best performance was achieved with GMMs, using five enrollment samples, and that the system did not loose much of the performance is the number of features reduced down to 15 (another relevant work can be seen in [6]). As the number of features is so low and each of them can even be coded in one single byte, the viability of integrating this modality with smart cards was a reality, and even the development of a match-on-card prototype was shown in [7].

$$d_{hamming}(x_i, t_i^m) = \#\{i \in \{1, ..., L\} / |x_i - t_i^m| > t_i^v\} \quad (3)$$

As already mentioned, it is of significant importance to show the success of the first commercial systems, because these systems demonstrated the viability of this biometric modality in real scenarios. The first unit shown was in 1972 from Identimat, but popularity was gained by the products of Recognition Systems. They developed their first prototype named HandKey ID3D before 1990, and improved such system in 1997 by launching HandKey II (shown in Fig. 1). Hundreds of thousand units have been sold, including applications in Universities, Airports, or Nuclear Plants. This technology has gained wide application and acceptance, especially in Access Control Systems, and in Time and Attendance Control.

Evolutions from Initial Works

From the results shown in the previous mentioned works, several R&D groups have worked in this biometric modality. They have improved the system in several ways. One of those working lines has been improving usability by removing the orientation pegs. Some researchers use a commercial scanner (e.g., [8, 9]), while others have worked not only in a peg-free, but also a contact-free system (e.g., [6, 10, 11]).

Other working lines have been focused in new feature extraction approaches. Some authors have increased the number of features, by including not only geometrical measurements, but also information

about the hand contour [12]. Kumar and Zhang [13] improved verification rates in 4–7%, by discretizazing features based on entropy studies. Gross et al. [9] have worked with Active Appearance Models. Others have worked in modeling hand contour and extracting features by curvature gradient (e.g., [14]). Ma et al. model hand geometry by using B-Spline curves [15]. Other authors work with neural networks (e.g., [12]), either for performing the whole identification process, or just for the comparison block.

Most of these studies claim error rates below 5%. Some authors give even better figures, approaching a 99% of identification accuracy. But even though, there are some major open issues regarding this biometric modality. One of those is the size of databases used for testing. Unfortunately in most works such databases are quite small, going up to 100 users with 10 photos per user.

There are still some open issues, especially nowadays and within some kind of population. The expansion in the use of jewelery, such as rings or piercings with a wide variety of shapes and sizes, can be considered as image artifacts by hand recognition systems, and lower the identification rates. Also tattoos, and specially those made in several colors can provoke the denial of use by the hand recognition system. These kind of problems have to be considered by new systems, to gain universality.

Usability and Multimodality

One of the most important facts related to this biometric modality is its great usability. It seems that users do not feel themselves afraid of using the system, neither of noting their privacy attacked. Kukula and Elliott [16] carried on a study that showed that 93% of users enjoyed the system, nearly all found it easy to use, and no one had privacy concerns. Kukula et al. [17] have also studied the effects of training and habituation in using the system, showing a better performance when users are familiar with the identification device.

This great usability, together with the fact that other biometric modalities use the same part of the body (e.g., palmprints or fingerprints), have pushed researchers based on multimodal biometrics to use this biometric modality. Fusion works using palmprints and hand geometry can be found in [18] or [8]. Other authors work even with three modalities, adding fingerprints to the previously mentioned ones, like in [19] or [9]. Or even some authors have developed multimodal prototypes with other non-hand-based modalities [20, 21].

Summary

Hand Geometry is a biometric modality whose promising features are the ease of use and high friendliness to the user. Furthermore, researchers have demonstrated that error rates below 5% are possible, and when applied to limited number of users, the level of performance is high enough for certain applications. Commercial products have found their business applications in Access Control Systems, as well as in Time and Attendance environments.

Related Entries

- ▶ Gaussian Mixture Models
- ▶ Hand Data Interchange Format
- ▶ Hand Databases and Evaluation
- ▶ Hand-Geometry Device
- ▶ Match-on-Card
- ▶ Multibiometrics
- ▶ User Acceptance

References

1. Schlage, R.S.: Main website. http://recognitionsystems.schlage.com/
2. Sidlauskas, D., Tamer, S.: In: Handbook of Biometrics, chap. Hand geometry recognition, pp. 91–107. Springer, Berlin (2008)
3. Golfarelli, M., Maio, D., Maltoni, D.: On the error-reject trade-off in biometric verification systems. IEEE Trans. Pattern Analy. Mach. Intell. **19**(7), 786–796 (1997)
4. Jain, A., Ross, A., Pankanti, S.: A prototype hand geometry-based verification system. In: Second International Conference on Audio and Video-Based Biometric Person Authentication (AVBPA), pp. 166–171 (1999)
5. Sanchez-Reillo, R., Sanchez-Avila, C., Gonzalez-Marcos, A.: Biometric identification through hand geometry measurements. Trans. Pattern Analy. Mach. Intell. **22**(10), 1168–1171 (2000). DOI 10.1109/34.879796
6. Kumar, A., Zhang, D.: Personal recognition using hand shape and texture. IEEE Trans on Image Process. **15**, 2454–2461 (2006)
7. Sanchez-Reillo, R., Gonzalez-Marcos, A.: Access control system with hand geometry verification and smart cards. IEEE Aerospace Electr. Syst. Mag. **15**(2), 45–48 (2000)

8. Savic, T., Pavesic, N.: Personal recognition based on an image of the palmar surface of the hand. Pattern Recogn. **40**(11), 3152–3163 (2007)
9. Ferrer, M., Morales, A., Travieso, C., Alonso, J.: Low cost multimodal biometric identification system based on hand geometry, palm and finger print texture. In: 2007 41st Annual IEEE International Carnahan Conference on Security Technology, pp. 52–58 (2007). DOI 10.1109/CCST.2007.4373467
10. Gross, R., Li, Y., Sweeney, L., Jiang, X., Xu, W., Yurovsky, D.: Robust hand geometry measurements for person identification using active appearance models. In: First IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007, pp. 1–6 (2007). DOI 10.1109/BTAS.2007.4401936
11. Zheng, G., Wang, C.J., Boult, T.E.: Application of projective invariants in hand geometry biometrics. IEEE Trans. Inform. Forens. Secur. 2(4), 758–768 (2007) DOI 10.1109/TIFS.2007.908239
12. Faundez-Zanuy, M.: Biometric verification of humans by means of hand geometry. In: 39th Annual 2005 International Carnahan Conference on Security Technology, CCST '05, pp. 61–67 (2005). DOI 10.1109/CCST.2005.1594816
13. Kumar, A., Zhang, D.: Hand-geometry recognition using entropy-based discretization. IEEE Trans. Inform. Forens. Secur. **2**(2), 181–187 (2007). DOI 10.1109/TIFS.2007.896915
14. Boreki, G., Zimmer, A.: Hand geometry: a new approach for feature extraction. In: Fourth IEEE Workshop on Automatic Identification Advanced Technologies, pp. 149–154 (2005). DOI 10.1109/AUTOID.2005.33
15. Ma, Y., Pollick, F., Hewitt, W.: Using b-spline curves for hand recognition. In: Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004, **3,** 274–277, Vol. 3 (2004). DOI 10.1109/ICPR.2004.1334520
16. Kukula, E., Elliott, S.: Implementation of hand geometry at purdue university's recreational center: an analysis of user perspectives and system performance. In: 39th Annual 2005 International Carnahan Conference on Security Technology, 2005, CCST '05, pp. 83–88 (2005). DOI 10.1109/CCST.2005.1594879
17. Kukula, E.P., Gresock, B.P., Elliott, S.J., Dunning, N.W.: Defining habituation using hand geometry. In: 2007 IEEE Workshop on Automatic Identification Advanced Technologies, pp. 242–246 (2007). DOI 10.1109/AUTOID.2007.380627
18. Kumar, A., Wong, D.C., Shen, H.C., Jain, A.K.: Personal authentication using hand images. Pattern Recogn. Lett. **27**(13), 1478–1486 (2006)
19. Yang, F., Ma, B., Wang, Q.X., Yao, D., Fang, C., Zhao, S., Zhou, X.: Information fusion of biometrics based-on fingerprint, hand-geometry and palm-print. In: 2007 IEEE Workshop on Automatic Identification Advanced Technologies, pp. 247–252 (2007). DOI 10.1109/AUTOID.2007.380628
20. Ross, A., Jain, A.: Information fusion in biometrics. Pattern Recogn. Lett. **24**(13), 2115–2125 (2003)
21. Jain, A., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. Pattern Recogn. **38**(12), 2270–2285 (2005)

Hand Geometry View Record – HGVR

Block of data that contains a hand silhouette captured from one camera point of view during one hand placement.

► Hand Data Interchange Format, Standardization

Hand Physiology

► Anatomy of Hand

Hand Shape

Nicolae Duta
Nuance Communications, Burlington, MA, USA

Synonym

Hand contour

Definition

A hand shape biometric system uses a camera or scanner-based device to acquire the hand image of a person from which shape information is extracted and compared against the information stored in a database to establish identity. Due to its limited discrimination power, a hand shape biometric system mostly operates in verification mode; that is the system *confirms or negates* the claimed identity of an individual.

Introduction

An increasing number of systems require positive identification before allowing an individual to use their services. Biometric systems are already employed in domains that require some sort of user verification. It is generally accepted that fingerprint and iris patterns can uniquely define each member of an extremely

large population which makes them suitable for large-scale recognition (establishing a subject's identity). However, in many small-population applications, because of privacy or limited resources, it is only needed to authenticate a person (confirm or deny the person's claimed identity). In these situations, traits with less discriminating power such as hand shape, hand geometry, voice or signature can be used.

As noted in [1], hand shape-based authentication is attractive due to the following reasons:

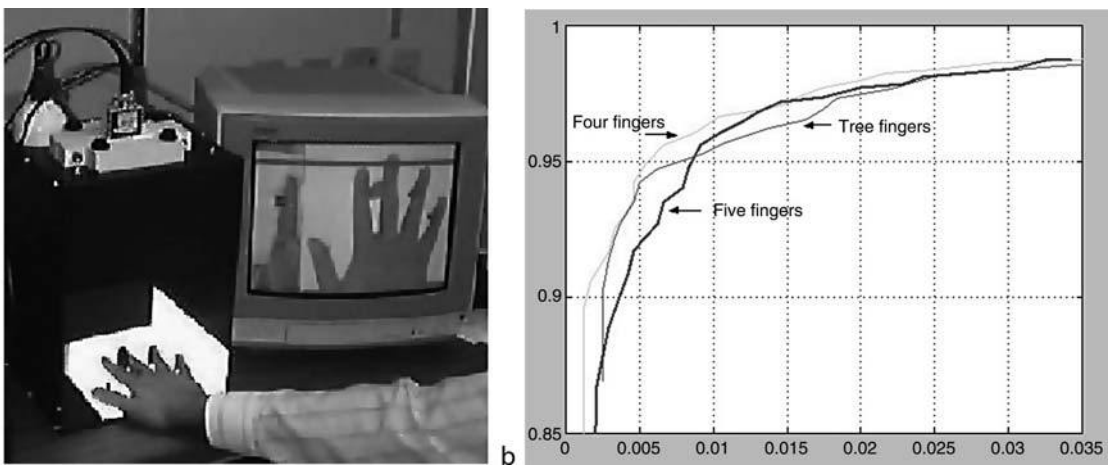
1. Hand shape can be captured in a relatively user convenient, nonintrusive manner by using inexpensive cameras.
2. Extracting the hand shape information requires only low resolution images, and the user templates can be efficiently stored (120-byte templates are reported in [1]).
3. This biometric modality is more acceptable to the public mainly because, it lacks criminal connotation.
4. Additional biometric features such as hand geometry, palmprints, and fingerprints can be easily integrated to an existing hand shape-based authentication system.

Operation of a Hand Shape-Based Biometric System

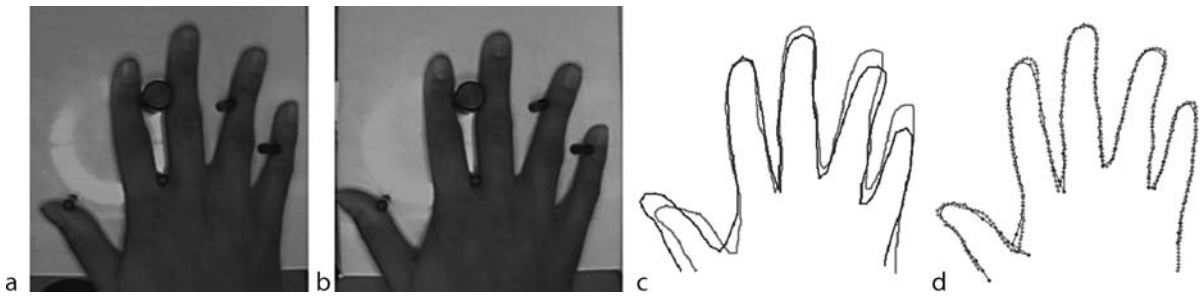
A hand shape-based biometric system operates according to the general diagram in [2] Fig. 2. In the enrollment stage, hand shape data is acquired from the registered users, feature sets are extracted from the acquired data, and one or multiple templates per individual are computed and stored in a database. In the deployment stage, one snapshot of the user's hand is captured; a feature set is computed and then compared to the user's templates in the database. Based on the comparison result, the claimed identity is accepted or denied. As described in [2], the system comprises the following modules: the sensor module, the feature extraction module, the matching module, and the decision-making module.

The sensor is usually a low/medium resolution CCD camera attached (beneath or above) to a platform on which the hand is placed (Fig. 1a). Some multimodal biometric systems capture the palm surface which includes both the hand shape and palmprints [3]. Other systems capture the dorsal surface of the hand from which only the hand contour can be extracted (Fig. 2, [1, 4, 5]). Some of the systems include on the platform 4–6 pegs to guide the placement of the user's hand [4, 5]. Several researchers noted that the guidance pegs deform the hand contour and decrease user convenience and proposed peg-less setups [1, 3, 6]. In a few systems, the sensor consisted of a 45 dots per inch (DPI) scanner [3, 6].

In the feature extraction module, a set of discriminating features is computed from a user's raw hand image(s). The hand images are first pre-processed in order to extract the hand contour and eliminate artifacts such as the guidance pegs, user rings, overlapping cuffs, or creases around the contour due to too light or



Hand Shape. Figure 1 (**a**) Example of a hand shape image acquisition system. (**b**) ROC curves for a hand shape-based verification system. The three curves correspond to feature vectors extracted from three, four and five fingers.



Hand Shape. **Figure 2** Hand shape alignment. Two scans of the same hand: (**a**–**b**) original images, (**c**) hand shapes extracted from (**a**) and (**b**) overlaid, and (**d**) finger aligned shapes (Mean alignment error = 2.20 pixels).

too heavy hand pressing. The pre-processing step can range from simple ▶ thresholding [1, 5] to sophisticate gray-level segmentation (▶ Image segmentation) [4]. Possible dents at the artifact location are smoothed by linear interpolation [4, 5] and/or morphologic operators [3].

In order to properly compare feature vectors extracted from hand images, one has to align the hand contours such that each feature is computed from the same region of the hand. Most of the older systems relied on the pegs to align the hand images. However, if the user is untrained or does not cooperate to proper use of the hand scanner, then the resulting images are not aligned (Fig. 2(c)) and the system's verification performance degrades [4]. Therefore, it is necessary to automatically align the acquired hand shapes before extracting the feature vectors used for verification. Due to the flexible nature of the palm and fingers, there may be no linear transformation which accurately aligns two hand contours. Hence, many of the proposed alignment procedures detect and align each finger separately. The simplest finger alignment method consists in registering the fingertip, the two adjacent valley points and several equally spaced points along the contour in between the three landmarks [5]. Similarly, a translation, rotation, and scaling can be found to align the finger with symmetry axis [6]. A more sophisticated alignment procedure (based on quasi-exhaustive polynomial search of point pair matching between two sets of contour points) is presented in [4]. This procedure has the advantage of always finding a good alignment even if the valley-point landmarks are not accurately detected. The alignment step can be avoided if the set of features extracted from the hand image is invariant to Euclidean transformations [1].

The hand shape can be modeled either explicitly as a set of 2D coordinates of several landmark points along the hand contour [4, 5] or implicitly as a binary image of the hand over an empty background [1, 6]. The two representations are intrinsically equivalent; each of them can be easily derived from the other. With both representations, dimensionality reduction procedures may have to be applied as the original data typically has a high dimensionality (see the fifth column in Table 1). The dimensionality reduction methods most used are ▶ principal component analysis (PCA) and independent component analysis (ICA), and are applied to either the original data or to a transformed version of the data (e.g., the Zernike moments in [1]).

One or several templates per user may be created during the enrollment stage and stored in the system's database. The templates are either the raw feature vectors computed from a user's hand images or the average of those feature vectors.

The matching module compares a user feature vector against the user's template(s) stored in the database in order to generate matching scores. Since the feature vectors are usually points in an N-dimensional Euclidean space, any metric distance can be used for computing a matching score: Euclidean distance [1], ▶ Mahalanobis distance, absolute (L_1) distance [6], correlation coefficient, etc. A few studies explicitly model the class-conditional probabilities under Gaussian assumptions [5]. As an exception, [3, 4] ▶ Procrustes shape distance can be used since the feature vectors are shapes corresponding to the hand contour. The matching score is a positive number which shows the dissimilarity between the user's hand and the templates in the database.

Hand Shape. Table 1 Comparison of some hand shape-based systems presented in the literature

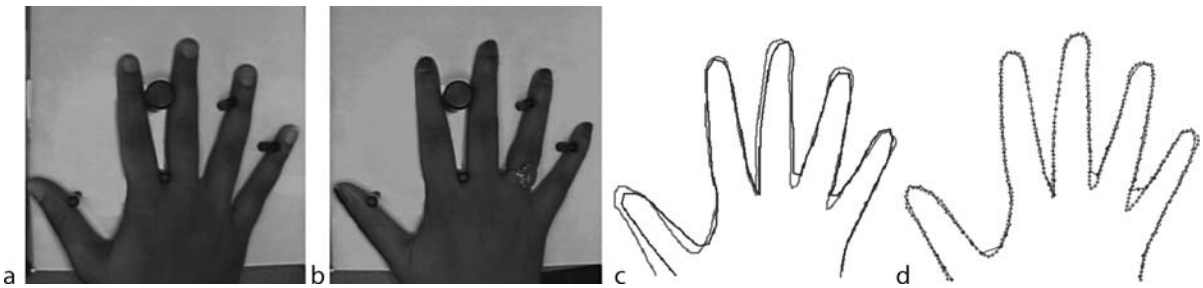
System	Population size	Samples/ person	Number/ type of templates	Features used	Similarity measure	Performance
[1]	40	10	5 (raw feature vectors)	Zernike moments of the binary hand image followed by PCA (30)	Euclidian	FAR = 0.01 FRR = 0.02 EER = 0.0164
[2]	53	2–15	1–14 (raw contours)	Hand contour coordinates (120–350 contour points)	Mean alignment error	FAR = 0.01 FRR = 0.06
[4]	51	10–20	1 (average) + multiple raw	Hand contour coordinates, angles (51–211 contour points)	Log-Likelihood under Gaussian assumption	EER = 0.00001 – 0.002
[5]	458	3	2 (raw feature vectors)	Contour coordinates (2048 points) ICA on binary hand image (458)	Modified Hausdorff L1, cosine distance	EER = 0.01– 0. 02

The final decision concerning the user's identity (identification) or the user's claimed identity (verification) is taken by the decision module. In verification mode, the decision is typically threshold based: if the matching score is below a given threshold the claimed identity is accepted, otherwise it is rejected. The threshold value is chosen based on the system's ROC curve such that the system satisfies some operating constraints (e.g., an upper bound on the false accept rate, an equal error rate, etc.). In identification mode, the incoming feature vector is typically assigned the identity of the closest database template if the distance to that closest template is lower than the verification threshold, otherwise the feature vector is considered to belong to an imposter.

Performance Evaluation

For most of the research systems, performance evaluation can only be based on the results and comparisons provided by their authors. Several enrollment and performance evaluation methodologies are discussed in [5]. If extensive enrollment data can be acquired, best performance is attained when an average template is computed from each user's enrollment measurements. However, such a system is less user-friendly and more difficult to deploy. A more realistic deployment scenario requires only one or few enrollment measurements per user. In such case, a template may actually be a raw feature vector and the main system parameter to estimate is the decision threshold. Most researchers split the available measurements for each user into an enrollment set and a testing set, and evaluate the threshold value based on the enrollment data. This has an advantage that the training data is representative for the test data, i.e., one expects to obtain an estimate of the decision threshold which works as well on the test data. In commercial deployments though, the system may be trained by the manufacturer while the enrollment is performed by the customer who has to use the factory set threshold.

It is difficult to directly compare the performance figures reported in the literature. The main reason is the absence of (1) a common benchmark data set and (2) standard enrollment and testing procedures. The different datasets used for research introduce several variation factors in the systems reported: (1) population size, (2) population age and/or structure, and (3) users' motivation to cooperate. Table 1 summarizes the performance of some research systems tested in identity verification mode. Some authors only report equal error rate (EER) figures while others include the system ROC curve. When the ROC curve was present, the estimated FRR rate corresponding to FAR = 0.01. As it can be seen in the last column of Table 1, most systems reported error rates on the order of 10^{-2}. A few systems have also been tested in identification mode and report identity recognition errors of 1–3% [6]. Note that no separate imposter population is used so the recognition error may be under estimated. Some authors integrate hand shape features with palm-based



Hand Shape. Figure 3 Hand shape alignment. Two scans of different hands: (**a**–**b**) original images, (**c**) hand shapes extracted from (**a**) and (**b**) overlaid, and (**d**) finger aligned shapes (Mean alignment error = 2.02 pixels).

features (which can be acquired through a single image measurement) and report verification error rates on the order of 10^{-3} [3].

Limitations of the Hand Shape-Based Biometric Systems

There are two factors which make a hand shape-based biometric system less accurate than a fingerprint or iris-based system [4]:

1. The hand shape is not unique within a large population. That is demonstrated in Fig. 3: after finger-alignment, the hand contours of two different users are almost identical. Therefore any geometric features extracted from the two aligned contours will be very similar and the system will likely confuse the identity of the two persons. In [4] three pairs of users with very similar hand shapes within a population of 53 persons were identified. However, the problem is alleviated if the system is used in verification mode since an imposter is less likely to know the identity of the registered user(s) whose hand shape best matches his or hers.
2. The human hand is a flexible object and its contour may suffer non-linear deformations when multiple hand images are acquired from the same person. That is demonstrated in Fig. 2 where a user's thumb appears to be longer in one of the images, a fact which makes the system reject its true identity. This problem is alleviated if the thumb (which can deform more than the other four fingers) is excluded from the feature vector calculation. Figure 1(b) compares the ROC curves corresponding to using all five fingers versus excluding the thumb and/or the little finger. The system which excludes the thumb exhibits a substantially better performance over the system which uses all five fingers.

Summary

Hand shape-based biometric systems have been successfully demonstrated for applications involving personal identity verification. Their ease of use, nonintrusiveness, public acceptance, integration capabilities, and small resource requirements have made hand shape popular among the different biometric modalities.

Related Entries

- ▶ Hand geometry
- ▶ Independent Component Analysis
- ▶ Multimodal systems
- ▶ Palmprint matching

References

1. Amayeh, G., Bebis, G., Erol, A., Nicolescu, M.: Peg-Free Hand Shape Verification Using High Order Zernike Moments. In: Proceedings of the IEEE Workshop on Biometrics in conjunction with CVPR06, New York, June, 2006
2. Ross, A., Jain, A.K.: Biometrics. Encyclopedia of Biometrics
3. Yörük, E., Dutagaci, H., Sankur, B.: Hand Biometrics. Image Vis. Comput. **24**, 483–497 (2006)
4. Jain, A.K., Duta, N.: Deformable Matching of Hand Shapes for User Verification. In: Proceedings of the IEEE International Conference on Image Processing (ICIP), Kobe, Japan, 857–861 (1999)
5. Veldhuis, R.N.J., Bazen, A.M., Booij, W., Hendrikse, A.J.: Hand-Geometry Recognition Based on Contour Parameters.

In: SPIE Biometric Technology for Human Identification II, Orlando, FL, 344–353 (2005)
6. Yörük, E., Konukoglu, E., Sankur, B., Darbon, J.: Shape-Based Hand Recognition. IEEE Trans. Image Process. **15**(7), 1803–1815 (2006)

Hand Shape Biometrics

► Hand Geometry

Hand Silhouette Data

► Hand Data Interchange Format, Standardization

Hand Structure

► Anatomy of Hand

Hand Vascular Recognition

► Back-of-Hand Vascular Recognition

Hand Veins

Graham Leedham
School of Information and Communication Griffith University, Queensland, Australia

Synonyms

Finger vein; Palm dorsal vein; Palm vein

Definition

In the human hand there is a complex structure of veins and blood vessels, many of which are just a few millimeters below the skin surface. Using noninvasive and safe imaging techniques it is possible to capture an image of the larger veins and blood vessels near the skin surface in various parts of the hand. These images are most readily obtained from the back of the hand and the palm of the hand. This vein structure, which is mostly invisible to the human eye, forms a pattern of interconnecting lines which is different from one individual to another and can be used as a physiological biometric. Two imaging methods can be used for safe, noninvasive imaging of veins near the skin surface: (1) Far infrared thermography, and (2) near infrared imaging. Far infrared imaging detects heat radiated from the hand and veins. Near infrared imaging detects infrared light reflected from a hand illuminated by near infrared light.

Introduction

In recent years, vein pattern biometrics has attracted increasing interest from research communities and industry. A system that scanned the back of a clenched fist to determine hand vein structure for verifying user identity was first reported in 1991 by Cambridge Consultants Ltd., in collaboration with the British Technology Group (BTG), who had been studying the hand vein pattern concept with the aim of developing a commercial system which they call Veincheck [1]. Though their product did not achieve much commercial success, the concept of hand vein patterns as a biometric was founded and has recently attracted further research and development interests to acquire the vein patterns in the back of the hand [2–5] and in the palm [6–8] as well as in the fingers [9].

A vein pattern is the vast network of blood vessels within a person's body carrying blood back to the heart. Anatomically, the distribution of veins in the body creates a vascular pattern which is believed to be distinct from person to person [10] and is also observed to be different between identical twins. The vascular patterns are reported to be stable over a long period of time, as a person's pattern of blood vessels is 'hardwired' into the body at birth, and remains relatively unaffected by aging, except for predictable

growth as seen in fingerprints. In addition, as the blood vessels are hidden underneath the skin and are invisible to the human eye, vein patterns are more difficult to copy or forge as compared to many other biometric features.

The properties of probable uniqueness, stability, and strong immunity to forgery of the vein pattern make it a potentially good physiological biometric to provide more secure and reliable person verification. However, as the vein patterns formed by superficial blood vessels lie underneath the skin surface, the invisibility of veins to simple visual inspection system creates significant difficulties in the acquisition of the vein pattern images. As the quality of the images plays a key role in all the subsequent processing stages of a vein pattern biometric system, the image acquisition is critical. In vein imaging for medical purposes, X-rays and ultrasonic scanning are used to obtain vascular images. While these methods can produce high quality images of blood vessels, X-ray imaging requires the invasive injection of a contrast agent into the blood stream and dosage of ionizing radiation which is dangerous with repeated exposure of even low level radiation. Ultrasonic imaging, while not known to have any adverse side effects, requires the application of a gel to the skin to improve the transmission of the sound waves as well as operator skill to obtain a good image. These constraints are not acceptable in general purpose biometric applications for security screening. Obtaining the vein pattern images quickly and accurately in a nonintrusive and noninvasive manner is a key challenge in the vein pattern biometric system.

Currently, the most effective means of obtaining images of veins near the surface of the skin without any invasive procedure or potentially dangerous side effect is to use the infrared range of the ▶ electromagnetic spectrum. Infrared imaging provides a contactless data acquisition method and requires no injection of any agents into the blood vessels. In the electromagnetic spectrum infrared refers to a specific region with wavelength typically spanning from 0.75 to 1,000 μm. This region is commonly further divided into four sub-bands: (1) Near infrared (0.75–2 μm); (2) Middle infrared (2–6 μm); (3) Far infrared (6–14 μm); (4) Extreme infrared (14–1,000 μm). Imaging objects within these four regions operates using different physical mechanisms and it results in images with significantly different properties. Far infrared and near infrared are the most suitable to capture images of human bodies.

Far Infrared Imaging

All objects emit infrared radiation when they are heated. The far infrared imaging technology forms an image passively using the infrared radiation emitted by the human body.

Principle of Far Infrared Imaging

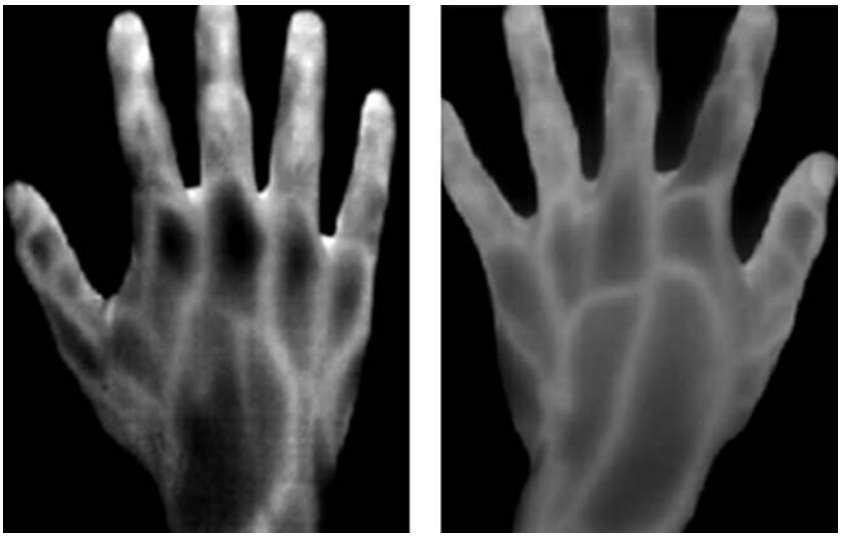
The total emissive power w is described by the Stefan-Boltzmann Law given in (1), where ξ is the emissivity of the object and $\sigma = 5{:}6703 \times 10^{-8} watt/m^2K^4$ is Stefan's constant. The relationship between the wavelength λ and black body temperature T is formulated by Wiens Displacement Law based on Planck's energy distribution law given in (2).

$$w = \xi \times \sigma \times T^4 \tag{1}$$

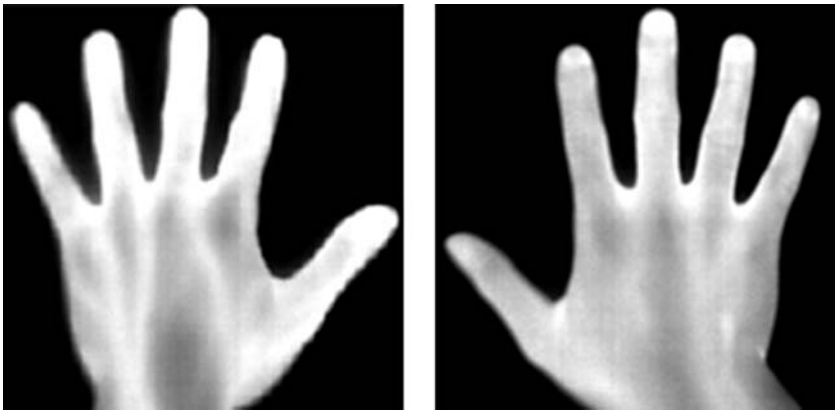
$$\lambda_{\max} = 2 : 9 \times 10^{-3}/T \tag{2}$$

Typically, a human body emits infrared radiation with wavelength in a range of 3–14 μm. These infrared waves radiate into the atmosphere and are attenuated according to the infrared transmittance spectrum of the atmosphere. At wavelengths of 3–5 μm and 8–14 μm, the radiant emittance of the infrared spectrum possesses the highest transmittance rate. Therefore, by using a thermal camera with detector sensitivity in the range of either 3–5 μm or of 8–14 μm, an image showing the heat distribution of the human body can be obtained. Medical researchers have observed that superficial veins have slightly higher temperature than the surrounding tissue. Therefore, using thermal imaging, an image of the heat distribution of body will display the location of veins just below the surface of the skin.

To acquire a far infrared image of the hand, the hand is usually placed on a flat surface with a far infrared camera focused on the hand from above. The far infrared camera will capture the temperature profile of the hand and transfer it to a computer which can convert the temperature data into either grey scale or color coding for display on a standard visual display and stored as a digital image for later computer processing.



Hand Veins. **Figure 1** Examples of far infrared images of hands at room temperature mapped to grey scale.



Hand Veins. **Figure 2** Examples of far infrared images of hands in a tropical climate mapped to grey scale.

Far-Infrared Vein Image Quality

Figures 1 and 2 show typical vein pattern images captured using a far infrared imaging method and converted to grey scale images. The darker the grey scale the cooler the pixel. The major vascular network in the back of the hand is successfully captured and appears as light grey lines as shown in Fig. 2. The images in Fig. 1 were captured in a normal office environment (approximately 20°C and 50% humidity) such that there is sufficient temperature difference at the skin surface to distinguish the location of the major veins beneath the skin. Figure 2 shows two images captured outdoors in a hot tropical climate (30–34°C and >80% humidity). In this case the temperature of the hand is closer to the blood temperature and there is insufficient difference in radiated thermal energy where the veins are located and the surrounding tissue to be discernable using the 0.08°C resolution of the camera used to capture this image. Far infrared imaging technology is very sensitive to external conditions which affect the temperature of the hand. In addition, far infrared imaging can only capture the major vein patterns. The smaller capillaries are not visible and the information contained in the large vein pattern is limited.

Near-Infrared Imaging

Human eyes can only see visible light which occupies a very narrow band (approximately 400–700 nm wavelength representing the color range from violet to red) of the entire electromagnetic spectrum. However, generally speaking, there is often more information contained in other bands of the electromagnetic spectrum reflected from the objects of interest. Some applications, such as remote sensing of crops, use special multispectral or hyperspectral imaging instruments to obtain the object images in a wide spread of bands of the electromagnetic spectrum. These images show more detail than is available in the visible light range. Similarly, while human vein patterns beneath the skin are invisible under normal visible light conditions they can be seen using near infrared imaging techniques.

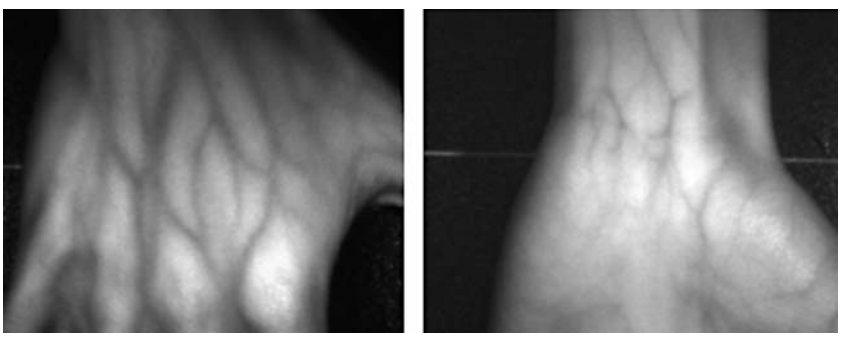
Principle of Near Infrared Imaging

Two special attributes of infrared radiation and human blood create a different principle for imaging vein patterns: (1) infrared radiation penetrates into biological tissue to a depth of approximately 3 mm, and (2) the reduced hemoglobin in venous blood absorbs more of the incident infrared radiation than the surrounding tissue [2]. Therefore, by shining an infrared light beam at the desired body part, an image can be captured using a CCD camera with an appropriate infrared filter attached to its lens. The location of veins within about 3 mm of the skin surface will appear as darker lines than the surrounding tissue because less infrared is reflected from where the veins are located due to it being absorbed in the blood. Biologically, the hemoglobin has the highest absorption of infrared light in the range of 800–900 nm [11]. Therefore, the wavelength of the infrared source should be selected to be within the near infrared region with wavelength around 800–900 nm. With this wavelength, it also avoids undesirable interference from the far infrared radiation (with a wavelength of 3–14 μm) emitted by the human body and the environment.

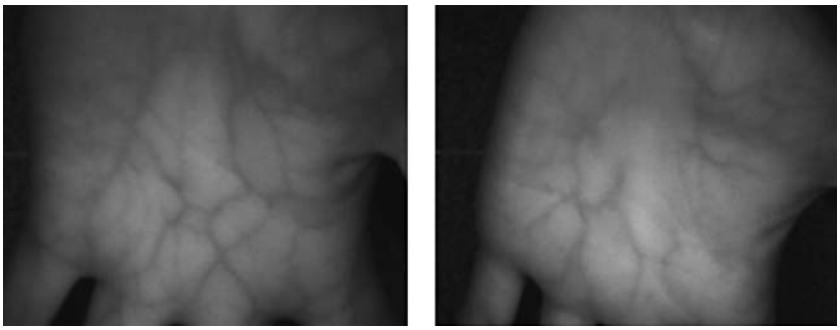
To acquire a near infrared image of the hand, the hand must be placed on a surface and evenly illuminated by infrared light. The infrared light should emit peak infrared radiation at about a wavelength of 850 nm. In order to obtain an image with this reflected infrared light from the hand, a CCD camera is needed whose spectral response also peaks at around 850 nm. Such cameras are readily available. To reduce the effect of visible light, an optical infrared filter of about 800 nm should be mounted on the cameras lens.

Near Infrared Vein Image Quality

Figures 3 and 4 show examples of vein pattern images captured using a near infrared camera. The veins just beneath the surface of the skin appear as dark lines. The near infrared imaging technique can capture the major vein patterns in the back of the hand as effectively as the far infra red imaging technique as shown Fig. 3. More importantly, the near infrared technique is capable of imaging some of the small veins lying in the palm and wrist areas. Unlike the image of the back of the hand, where only major veins are visible, the vein pattern in the palm is far more complex and potentially contains more information than an image of the back of the hand. This is important because it significantly increases the discrimination power of the vein pattern biometrics when the size of user group is



Hand Veins. Figure 3 Examples of near infrared images of the back of the hand and the underside of the wrist.



Hand Veins. **Figure 4** Examples of near infrared image of the palm of the hand.

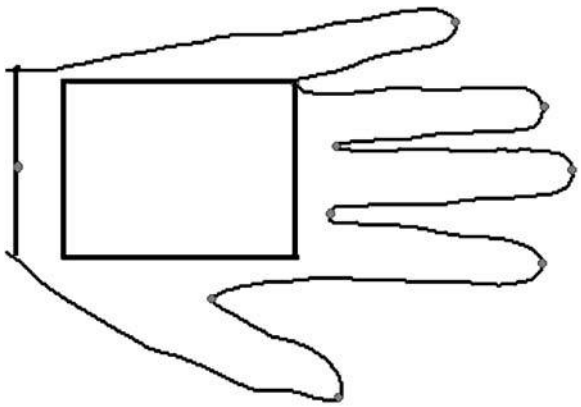
large. Near infrared imaging technique is more tolerant to the external environment and the subject's body temperature than far infrared imaging. However, near infrared images of vein patterns suffers from the disadvantage that visible marks on the skin surface are also visible in the image, which can corrupt the structure of the vein patterns and lead to problems in the later imaging processing and pattern recognition stages. The palm lines are also visible with the vein patterns as seen in Fig. 4. While human beings are capable of distinguishing these lines from the vein patterns in the image, it requires extra effort to remove these defects using automatic processing of these images and is particularly difficult if, for example, there are marks on the palm which are similar in appearance to veins, and for example when the person has drawn lines on the hand using a black pen.

A Hand Vein Pattern Matching System

A hand vein recognition system will typically consist of the following processing stages:

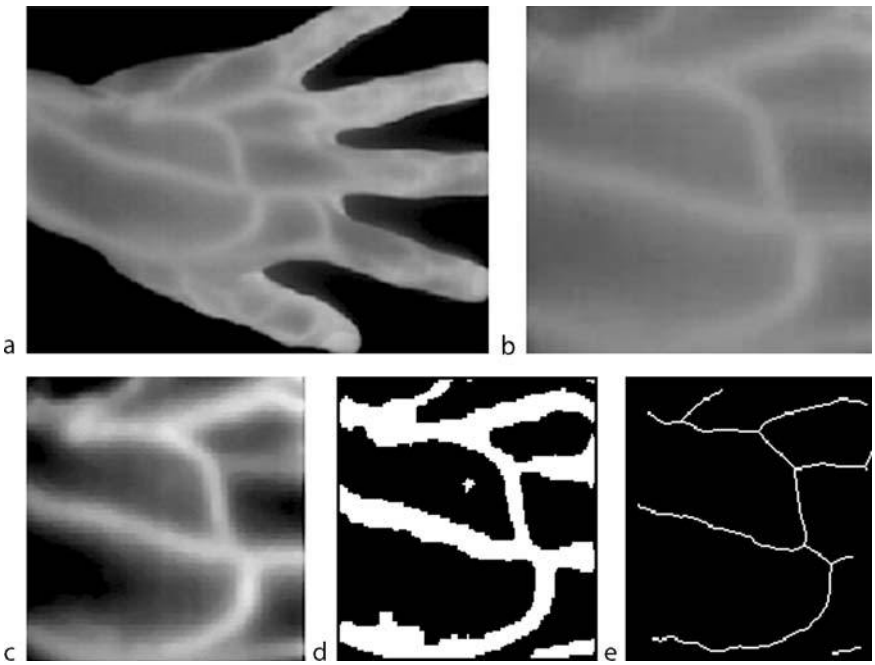
1. Hand Vein Image Acquisition
2. Region of Interest Location and Image Enhancement
3. Vein Pattern Extraction
4. Feature Extraction and Matching Against a Database of Vein Patterns
5. Decision

The methods and problems associated with Stage 1 – obtaining an image of the veins in the hand have been described above. Stage 2 – locating the region of interest and enhancing the image – is usually achieved by extracting the profile of the hand and locating the valleys between the fingers and the thumb, and using these as reference points to identify a region (usually a rectangle) on the palm or back of the hand as shown in Fig. 5. Image enhancement is needed because the clearness of the vein pattern in the extracted region of interest varies from image to image, therefore, the quality of these images need to be enhanced before further processing. There are many pre-processing techniques available in image processing for image enhancement. The choice of which to use will depend on the quality and nature of the image. An example of region of interest location, extraction, and image enhancement for a far infrared grey scale image of the back of a hand is shown in Fig. 6(a-c).



Hand Veins. **Figure 5** The typical region of interest in palm of back of hand vein imaging.

After image enhancement, processing in Stage 3 is required to extract the vein patterns in the region of interest. This involves separation of the vein pattern from the background. Due to the fact that the grey-level intensity values of the vein vary at different



Hand Veins. Figure 6 Example of (**a**) Original far infrared image of the back of the hand, (**b**) Extraction of the region of interest, (**c**) Image enhancement of the region of interest, (**d**) Extraction of the vein lines, (**e**) Skeltonization of the vein lines.

locations in the image, ► global thresholding techniques do not usually provide satisfactory results. Hence, a ► local adaptive thresholding algorithm is usually required to separate the vein patterns from the background. The binary image in Fig. 6(d) shows a typical result of the vein image after thresholding. The vein image may be processed in this binary form or the shape of the vein pattern is extracted as a skeleton, or one pixel wide line image, of the vein path. The result of a typical skeletonization process is shown in Fig. 6(e).

Stage 4 of the recognition process involves the extraction of features from the vein pattern and matching these against the same features extracted from reference patterns collected from known individuals and stored in a database of template or reference vein patterns. This stage remains an area for further research. The features can be extracted from the grey scale image (Fig. 6(c)), the binary image (Fig. 6(d)) or the skeletonized image (Fig. 6(e)). Previous research has investigated the matching of vein patterns using the Hausdorff distance as used in face recognition [12] and the extraction of minutiae from the skeletonized image in a similar manner to that frequently applied to fingerprint images [13].

The major factor restricting further investigation of hand veins as a biometric is the lack of a large database of hand vein images for research study. All reported work carried out to date has involved relatively small databases collected by the individual researchers. It is therefore not possible to compare performance results or predict the likely false acceptance and false rejection rates that might be expected of hand vein biometrics.

Summary

The study of hand veins as a biometric has been investigated sporadically since about 1990. The most successful imaging methods use near and far infrared imaging. Far infrared imaging can capture images of the large veins in the back of the hand but has difficulties in capturing vein images in the palm because of the relatively small size of the veins and the resulting small amount of thermal energy they radiate. Far infrared imaging is very sensitive to ambient temperature and varying human body temperature, which can be significant in an extremity such as the hand. Near infrared imaging produces good quality images of

veins just below the surface of the skin as observed when capturing vein patterns in the back of the hand or the palm. Near infrared is more tolerant to environmental changes because the technique measures reflected infrared and not transmitted infrared. The major problem with near infrared images is the retention of visual features such as marks on the skin and hairs. Detailed study of the processing and matching of hand vein images needs to be carried out to fully assess the potential of hand veins as a biometric.

Related Entries

- ▶ Anatomy of Hand
- ▶ Finger Geometry, 3D
- ▶ Hand Databases and Evaluation
- ▶ Hand Geometry
- ▶ Hand Shape
- ▶ Palm Vein

References

1. MacGregor, P., Welford, R.: Veincheck: Imaging for security and personnel identification. Adv. Imaging **6**(7), 52–56 (1991)
2. Cross, J.M., Smith, C.L.: Thermographic imaging of subcutaneous vascular network of the back of the hand for biometric identification. In: Proceedings of IEEE 29th International Carnahan Conference on Security Technology. Sanderstead, IEEE Publisher, Surrey, UK (1995)
3. Im, S.K., Park, H.M., Kim, S.W., Chung, C.K., Choi, H.S.: Improved vein pattern extracting algorithm and its implementation, In: Digest of technical papers of International Conference on Consumer Electronics, pp. 2–3, IEEE Publisher, Los Angeles, USA (2000)
4. Wang, L., Leedham, C.G.: A thermal hand vein pattern verification system. In: Proceedings of International Conference on Advances in Pattern Recognition, pp. 58–65, Springer, Bath, UK (2005)
5. Kumar, A., Prathyusha, K.V.: Personal authentication using hand vein triangulation. In: Proceedings of the SPIE, Biometric Technology for Human Identification V., vol. 6944, pp. 69440E–69440E-13 (2008)
6. Lin, C.L., Fan, K.C.: Biometric verification using thermal images of palm-dorsa vein patterns, IEEE Transactions Circuits and Systems for Video Technology **14**(2), 199–213 (2004)
7. Fujitsu-Laboratories-Ltd., Fujitsu laboratories develops technology for world's first contactless palm vein pattern biometric authentication system. Available at http://pr.fujitsu.com/en/news/2003/03/31.html (2003)
8. Wang, J.-G., Yau W.-Y., Suwandy, A., Sung E.: Person recognition by fusing palm vein images based on Laplacianpalm representation. Pattern Recognit. **41**(5), 1514–1527 (2008)
9. Miura, N., Nagasaka, A., Miyatake, T.: Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. Mach. Vis. Appl. **15**, 194–203 (2004)
10. Jain, A., Bolle, R.M., Pankanti, S.: Biometrics: personal identification in networked society. Kluwer Academic Publishers, Dordrecht (1999)
11. Fujimas, I., Nakazawa, H.: Structural and functional tissue analysis under skin using near infra-red spectral imaging. In: Proceedings of the First Joint BMES/EMBS Conference **2**, 1114 (1999)
12. Wang, L., Leedham, G.: Infra-red imaging of hand vein patterns for biometric identification purposes. IET Proceedings on Vision **1**(3–4), 113–122 (2007)
13. Wang, L., Leedham, G., Cho, S.Y.: Minutiae feature analysis for infra-red hand-vein pattern biometrics. Pattern Recognit. **41**(3), 920–929 (2008)

H

Hand Vein Identification

▶ Back-of-Hand Vascular Recognition

Hand Vein Verification

▶ Back-of-Hand Vascular Recognition

Hand-Geometry Device

Vitomir Štruc, Nikola Pavešić
Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, SI-1000 Ljubljana, Slovenia

Synonyms

Hand-geometry reader; Hand-geometry scanner; Hand-scanning device

Definition

Hand-geometry devices are specially designed biometric devices used for capturing the geometric

characteristics (e.g., the length, width, thickness and curvature of the fingers, the palm size, and the distances between joints) of a human hand for hand-geometry-based identity verification. A typical hand-geometry device records images of the lateral and dorsal parts of the hand using a charge-coupled device (CCD) camera that is mounted above a flat surface on which the person presented to the device places his/her hand. The set of geometrical features extracted from these images is then matched against a prerecorded template stored in the device's database. Depending on the result of this matching procedure, the identity of the person presented to the device is either verified or not.

Introduction

Hand-geometry devices are among the earliest commercially available biometric devices for automated identity verification [1]. The production and distribution of the first commercial hand-geometry device, called the Identimat, was launched in the early 1970s by the Identimation Corp., which adopted a hand-reader concept developed and patented by Robert P. Miller [2]. Like with Miller's original design, Identimation's device used spatial characteristics of the index, middle, ring, and little fingers as the means of establishing the identity of a person. It utilized a number of electromechanical components and photoelectric cells to measure the length of the four digits and compare them to finger-length measurements that were previously recorded and stored on an identification card. The device was very well received on the market and was eventually installed for access-control purposes at several high-security facilities run, for example, by the U.S. Department of Energy, Western Electric, and U.S. Naval Intelligence [3].

Encouraged by the success of Identimation's hand-based verification system and the growing demand for reliable and user-friendly verification schemes, several other companies tried to enter the hand-biometry market during the 1970s and early 1980s. They developed numerous prototypes, using ideas and device designs from early patents (e.g., [2, 4, 5]); however, most of them never actually made it to the market. One of the few exceptions was the "3D hand profile identification apparatus" devised by David Sidlauskas [6]. His device featured a ▶ platen on which a person placed his/her hand and a digital camera that captured images of the hand's side and top views. Discriminative hand characteristics extracted from these images were then employed for the identity verification. Unlike previously developed hand-geometry readers, Sidlauskas' device did not rely solely on two-dimensional measurements of the hand, but used ▶ orthographic scanning to capture the hand's three-dimensional structure. Later manufactured under the commercial name ID3D by Recognition Systems Inc. (RSI) [7], it became an important milestone in the field of hand-geometry-based verification and is, albeit in a much refined form, still on the market today.

By the late 1990s, manufacturers of commercial hand-geometry devices (such as RSI) were the only driving force in the development of hand-geometry-based biometric technology. However, in the past decade, due to advancements in the fields of biometrics and computer vision, the academic community has taken a more active role in the development of hand-geometry devices, e.g., [1, 8, 9, 10].

Description of the Device

Unlike early hand-geometry devices (e.g., [2, 4]), which were primarily based on electro-mechanical components, modern devices, such as the one presented in Fig. 1, use imaging technology and internal software to capture and process the images of a person's hand and to extract the geometrical features, e.g., the lengths, widths, thicknesses, and curvatures of the fingers, and the width, thickness, and area of the palm, that are used for the identity verification.

Hand-Geometry Device. **Figure 1** A commercial hand-geometry device [7].

The basic design and use of a hand-geometry device is quite simple. When a person places his/her hand on the device's reflective, flat surface, referred to as the platen, he/she first has to align his/her fingers with a number of guiding pegs that direct the hand to a predefined position. The pegs are equipped with pressure sensors which, when enough pressure is applied to them, simultaneously trigger the charge-coupled device (CCD) camera (commercial devices typically use a 32,000 pixel CCD camera) and the infrared light source (e.g., light emitting diodes) positioned above the device's platen. The platen then reflects the emitted light back to the camera and an image is recorded. However, as parts of the platen are covered with the person's hand, some of the infrared light is absorbed and only a silhouette is visible in the resulting image [11, 12, 13]. Because of the design of modern hand-geometry devices, which feature a side-mounted mirror inclined at 45° to the platen, the acquired silhouette image contains both the shape of the dorsal (i.e, the top view) as well as the lateral (i.e., the side view) surfaces of the hand [9]. Once recorded, internal software extracts a number (more than 90 in commercial devices) of geometrical features from the silhouette image and uses them to verify the identity of the person presented to the device.

However, before a person can use the device, he/she first has to enroll. During the enrollment phase, the device captures several images of the person's hand, extracts geometrical features from each of these images and uses them for the calculation of the template. The template is then stored in the memory of the device or on an identification card (i.e., a ▶ smart card) and is later retrieved for comparison. A similar procedure is required when a person presented to the device is trying to verify his/her identity. First, the person claims an identity by entering a personal identification number (PIN) or by swiping an identification card (depending on the input mechanism provided by the device at hand) through a card-reader module connected to the device. The device then proceeds with the image-acquisition and feature-extraction stages and finally recalls (either from an internal memory or from the smart card) the template associated with the claimed identity for comparison. In the final step, a matching procedure is applied to decide whether or not the person presented to the device is who he/she claims to be [7, 9, 11, 12]. In the case of a positive decision, i.e., the identity of the person is verified, the device usually updates the template to account for possible changes in the geometry of the person's hand (which is especially important when the device is used by children, whose hand-geometry is changing fast) and stores a new template for future verification attempts in the device's memory (or ID card). This process is commonly referred to as template averaging [11].

While typical hand-geometry devices are designed to be used in conjunction with the right hand, it is possible for a person to enroll and verify his/her identity using the left hand. In this case, the (left) hand is placed on the platen with the palm facing upwards [11]. As only the geometry of the hand is of significance, this has no negative effect on the verification accuracy of the device.

There are also commercial devices available on the market that do not use the geometry of the whole hand to verify the identity of a person, but accomplish this task based on measurements of only two fingers.

The main part of the device is a camera-based sensor that uses three-dimensional scanning technology to capture the structure of the index and middle finger (of either hand) of the person presented to the device. From these scans, a set of geometrical features is extracted and matched against a template recorded during the enrollment session. Depending on the outcome of the matching procedure, the identity of the person presented to the device is either verified or not [14].

Research Trends

In recent years, many research groups from private companies as well as academic institutions have directed their research towards hand-geometry-based identity verification. They are developing new verification schemes that require new kinds of hand-geometry devices, different from those available on the market today. The main trend at present is to design devices that require no pegs to control the placement of the hand. These designs, like the current commercial devices, still feature a platen upon which the person places his/her hand. However, as there is no guiding mechanism, the hand is simply positioned on the platen with the fingers spread naturally. A CCD camera or a digital scanner then captures images of the hand from which pose-independent geometrical features are extracted and used to verify the identity of the person presented to the device [10]. Peg-free designs are

commonly considered to have a number of advantages over classical hand-geometry devices: first, they do not cause any deformations of the hand's silhouette shape as no contact with the guiding pegs occurs; second, they reduce the chances of the person incorrectly positioning his/her hand; and third, they allow people with small hands, e.g., children, who might have problems reaching all the pressure sensors of current commercial devices and so are unable to initiate the verification procedure, to use the device. Although numerous experimental designs following the described trend were presented in the literature, none of them has yet succeeded in passing the prototype stage [9, 10].

Characteristics

Over the past 30 years, hand-geometry-based recognition has become one of the most popular biometric technologies for physical access control and time-and-attendance applications. The broad success of hand-geometry devices in these specific application areas was triggered by various human and operational factors, among which the following are the most important [9]:

1. *User acceptance*. Hand-geometry devices offer a fast, easy to use, and fairly reliable method of user authentication, and are therefore enjoying a relatively high level of public acceptance. A survey conducted by the Sandia National Laboratories [15] in 1991 reported that most of the test subjects favored hand-geometry devices over other devices based on fingerprint, signature, retina, or face biometrics.
2. *Functionality*. Hand-geometry devices can operate in harsh environmental conditions and are therefore suitable for indoor as well as outdoor deployment [10]. Furthermore, as they rely only on the geometric structure of the hand, while ignoring its surface details, they are to some extent insensitive to the presence of dirt or dust, which makes them a preferable choice for access control and time-and-attendance applications in labor-intensive branches such as the construction industry [7].
3. *Template size*. The memory requirements for storing a template are the lowest among all biometric technologies. With a size of 9 bytes (or 20 bytes for devices based on the geometry of just two fingers) they are significantly lower than those imposed by other modalities [11, 12]. Such a small size is advantageous for three reasons: first, when a hand-geometry device is operating as a stand-alone unit, it allows a large number of templates to be stored in the device's internal memory; second, it saves processing time; and third, it permits the storage of user-templates on identification cards.
4. *Failure to enroll (FTE) and failure to acquire (FTA) rates*. Hand-geometry devices can be used by most of the world's population, except for some individuals who suffer from severe arthritic conditions and are therefore unable to correctly position their hands on the device's platen. For this reason, hand-geometry devices exhibit relatively low FTE and FTA rates, when compared to other biometric devices. A recent study involving 200 participants showed [16] that the FTE and FTA rates for hand-geometry devices were the lowest among all the tested devices (tested were face, iris, vein, voice, hand-geometry, and fingerprint scanners).

Although the presented characteristics resulted in the widespread use of hand-geometry devices for access control and time-and-attendance purposes, there are, nevertheless, still a number of shortcomings that limit their use in other application areas, the most significant being:

1. *Size*. Typical hand-geometry devices are designed to accommodate the whole human hand (or at least two fingers) and are consequently significantly larger than other devices used for capturing the biometric traits of a person (e.g., face, voice, and fingerprints). This fact makes them unsuitable for security applications, where a compact size for the biometric device is preferable (e.g., laptops and mobile devices) [12].
2. *Cost*. Commercial hand-geometry devices can cost considerably more than, for example, fingerprint scanners, which target a similar market segment (i.e., access control and time-and-attendance). With a price of approximately $1,000–2,000, they are among the more expensive biometric technologies [9, 12].
3. *Performance*. The false-acceptance (FA) and false-rejection (FR) rates for hand-geometry-based security systems are typically higher than those of fingerprint-, palmprint- or iris-based systems, which makes hand-geometry devices suitable only for low/medium security applications. Several

independent studies (e.g., [13, 15, 16]) reported that commercial hand-geometry devices achieve FA and FR rates in the range of 0.1–1.0% at the equal-error operating point.

Operation

Commercial hand-geometry devices are capable of operating in two distinct types of configurations: as stand-alone units or as part of a networked system [7, 11, 17].

1. *Stand-alone units.* While hand-geometry devices deployed for time-and-attendance monitoring typically require an additional time clock and a computer to record and retrieve information about the arrival and departure times of people, they can, nevertheless, be used for access control purposes without the need for any additional components. A hand-geometry device can, for example, directly control the locking mechanism of a door and release it if the identity of the person trying to gain access to the secured facilities is successfully verified. In this stand-alone configuration, the devices are suitable only as access control systems for single doors (e.g., main entrances, doors to sensitive areas such as computer rooms, storage areas, etc.), while a (networked) system of hand-geometry devices is needed when multiple doors have to be secured. The number of people that can be enrolled in a stand-alone device is limited by the storage capabilities of the device, as all user templates are stored locally in the device's internal memory. Furthermore, as no other means are available, for example, a central computer, all administrative tasks have to be preformed with the help of the device's keypad [7, 11, 17].
2. *Networked system.* Commercial devices support a number of communication standards and protocols (e.g., RS-485, RS-422, RS-232, and TCP/IP) that can be used to connect an arbitrary number of devices with a host computer to form a networked system. In contrast to stand-alone units, networked systems are commonly employed in applications that require multiple hand-geometry devices (e.g., access control to facilities with several entrances). While these requirements can be met with several stand-alone devices, the use of a networked system has a number of advantages: first, a person does not have to undergo the inconvenience of enrolling at each of the units, but is able to enroll at a single location and retrieve his/her template for comparison from a central storage location at any unit (for which he/she has access rights) of the network; second, all door activities and time (and/or attendance) records can be stored and viewed via a central computer, making system monitoring simple and efficient; and third, a networked system enables the centralized management of user profiles (e.g., their access rights, and deletions) [7, 17].

Commercial devices are also able to emulate standard card-reader units, which makes it easy to integrate them into existing security systems. When employed in the card-reader emulation mode, the hand-geometry device, upon successful verification, simply forwards the user's identification number in an appropriate format to the card-reader module, which then proceeds as if the identification number had been read from an identification card. Several card protocols are commonly supported, the main ones being Wiegand, barcode, and magnetic stripe [7, 11, 17]. In fact, there are two standards defining the data-interchange formats of hand-geometry devices: the "ANSI INCITS 396-2005 Hand Geometry Interchange format" and its international counterpart the "ISO/IEC 19794-10:2007 Biometric Data Interchange Format – Part 10: Hand Geometry silhouette data." The standards define both the format and the content for the exchange of the hand-silhouette data, and are aimed at increasing the interoperability of hand-geometry devices [18].

Summary

Among the different biometric devices available on the market, hand-geometry devices have emerged as the preferred choice for physical access control and time-and-attendance applications, especially in harsh environments where other devices might have problems in reliably verifying the identity of a person. They are based on a field-proven technology that by today has been in use for more than 20 years. However, research is already on the way to produce the next generation of hand-geometry devices, which will undoubtedly result in smaller, faster, and more user-friendly units. Several

research groups have already developed prototypes that require no guiding pegs to capture an image suitable for the extraction of hand-geometry features. While these prototypes still need time to mature, they are a clear indication of future trends in the development of hand-geometry devices.

Related Entries

- ▶ Anatomy of Hand
- ▶ Biometric Sensor and Device, Overview
- ▶ Hand Recognition

References

1. Jain, A.K., Ross, A., Pankanti, S.: A prototype hand geometry-based verification system. In: Proceedings of the Second International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), Washington D.C., USA, pp. 166–171 (1999)
2. Miller, R.P.: Finger dimension comparison identification system. US Patent No. 3576538 (1971)
3. Miller, B.L.: Vital Signs of Identity. IEEE Spectr **31**(2), 22–30 (1994)
4. Ernst, R.H.: Hand ID system. US Patent No. 3576537 (1971)
5. Jacoby, I.H., Giordano, A.J., Fioretti, W.H.: Personnel Identification Apparatus. US Patent No. 3648240 (1972)
6. Sidlauskas, D.P.: 3D hand profile identification apparatus. US Patent No. 4736203 (1988)
7. Homepage of Recognition Systems Inc.: http://www.recogsys.com
8. Sanches-Reillo, R., Sanches-Avila, C., Gonzales-Marcos, A.: Biometric identification through hand geometry measurements. IEEE Trans. Pattern Anal. Mach. Intell. **22**(10), 1168–1171 (2000)
9. Pavešić, N., Ribarić, S., Ribarić, D.: Personal authentication using hand-geometry and palmprint features – the state of the art. In: Proceedings of the Workshop: Biometrics: Challenges Arising from Theory to Practice, Cambridge, UK, pp. 17–26 (2004)
10. Dutagaci, H., Sankur, B., Yoruk, E.: A comparative analysis of global hand appearance based person recognition. J. Electron. Imaging, **17**(1), 011018 (2008)
11. Zunkel, R.: Hand geometry based verification. In: Jain, A.R., Bolle, R., Pankanti, S. (eds.) Biometrics: Personal Identification in Network Society. Kluwer, Dordecht (1999)
12. Homepage of International Biometric Group: http://www.biometricgroup.com
13. Kukula, E., Elliott, S.: Implementation of Hand Geometry: An Analysis of User Perspectives and System Performance. IEEE Aerospace and Electronic Systems Magazine **21**(3), 3–9 (2006)
14. Homepage of Biomet Partners Inc.: http://www.biomet.ch/
15. Holmes, J.P., Wright, L.J., Maxwell, R.L.: A Performance Evaluation of Biometric Identification Devices. Sandia National Laboratories, Report, U.S.A. (1991)
16. Mansfield, T., Kelly, G., Chandler, D., Kane, J.: Biometric Product Testing: Final Report. Technical Report CESG Contract X92A/4009309, Centre for Mathematics and Scientific Computing, National Physics Laboratory, Middlesex, UK (2001)
17. Spence, B.: Biometrics in Physical Access Control: Issues, Status and Trends. Available at: http://www.recogsys.com
18. Homepage of the American National Standard Institute: http://www.ansi.org/

Hand-Geometry Reader

▶ Hand-Geometry Device

Hand-Geometry Scanner

▶ Hand-Geometry Device

Hand-Held Devices

A hand-held device is a pocket-sized computing device, typically, comprising of a small visual display screen for user output and a miniature keyboard or touch screen for user input. New hand-held devices include a number of sensors that can be used to acquire biometric data, e.g., touch-screens (signature and handwriting), fingerprint sensors, microphones (speech), cameras (face, video), etc.

▶ Fingerprint Databases and Evaluation

Handprint

In a handprint writing style, the writer is inclined to write each individual character in an isolated fashion. In other words, there is no connection between adjacent letters.

▶ Signature Sample Synthesis

Handprint Sensor

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Hand-Scanning Device

► Hand-Geometry Device

Handwriting Sample Synthesis

► Signature Sample Synthesis

Handwriting Synthesis

► Signature Sample Synthesis

Handwritten Signature Recognition

► Signature Recognition

Head Pose Analysis

► Face Pose Analysis

Head Yaw/Tilt/Roll

Yaw, Pitch, and Roll are the three angles of rotation that describe changes in the orientation of a 3D object, which in the case of face pose are with respect to the face being in upright position facing the camera/sensor. Yaw is the left/right rotation angle (head turning), pitch is the up/down rotation angle (head nodding) and roll is the sideways rotation angle (head tilting while facing the camera/sensor).

► Face Pose Analysis

Headspace

Headspace is the gaseous phase above a sample (liquid or solid) containing the volatile and semi-volatile compounds released by the substance.

► Odor Biometrics

Helper Data

Information extracted from biometric data to assist aligning or retrieving original biometric template. It should not leak any information about original template but assure alignment accuracy to some extent.

► User Interface, System Design

Heterogeneous

Heterogeneous is an adjective used to describe something that has a large amount of variants or different forms. Heterogeneity (a noun), resulting from having different properties, leads to a difference in expected results, more than can be accounted for by chance.

► Heterogeneous Face Biometrics

Heterogeneous Face Biometrics

Stan Z. Li
Biometrics and Security Research & National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, Beijing, China

Synonyms

Cross-modality face biometrics; Heterogenous face image matching

Definition

Face images can be captured in different spectral bands, *e.g.*, Visual (VIS), near infrared (NIR), or thermal infrared (TIR), or as measurements of 3D facial shape. These different image types, due to different image formation characteristics, are said to be ► heterogeneous. More generally, even within the VIS type, the face images can come from different image sensors, such as charge coupled device (CCD) and complementary metal oxide semiconductor (CMOS) cameras, photo scans, face sketches, under different illumination conditions, with different image resolutions and different image quality. These are heterogeneities in the broad sense. Although heterogeneous face images of a given person differ by pixel values, the identity of the face should be classified as the same. The processing and matching of these diverse face images is collectively referred to as ► heterogeneous face biometrics (HFBs).

Introduction

Different types of face biometrics have been developed, including visual (VIS) (see a survey in [1]), near infrared (NIR) [2], thermal infrared (TIR) [3], and 3D [4] image based. In each case, it is assumed that both the enrollment and query face images are of the same type. This assumption results in the *homogeneous* processing and matching of face images. Two different image types are said to be heterogeneous if they have different image formation characteristics. Although heterogeneous face images of a given person may differ significantly by pixel values, the identity of the face should be classified as the same notwithstanding the source, illumination, and quality of the image. The processing and matching of these diverse face images is collectively referred to as heterogeneous face biometrics (HFBs).

HFBs provide engines and systems for matching across different imaging systems, either in different spectral bands or modalities. These are considered HFBs (in the *true sense*). These involve comparison of face images between VIS, NIR, TIR, and 3D face images. Heterogeneities in the same type of imaging system are, in the *broad sense*, such as variations dealt with by the conventional VIS face recognition, including photo scans and face sketches. Arguably, the true sense HFBs are more challenging than the broad sense HFBs of VIS-VIS, NIR-NIR, and 3D-3D comparisons, because the former contain more heterogeneous factors and larger variations than the latter.

Recent developments have led to several proposals of HFBs that include matching between VIS and face sketch [5], VIS and NIR [6], and 3D and NIR [7], and also the reconstruction of the facial shape from an NIR image [8]. The MBGC (Multiple Biometric Grand Challenge) tests organized by NIST (National Institute of Standards and Technology) has devised the scenario of matching between VIS and NIR face images as one of its experiments [9], with the purpose of examining the feasibility of NIR-VIS face matching and determining how its fusion with VIS-VIS face biometric could improve the overall performance. In the following sections, the heterogeneities encountered in HFBs are discussed and related research issues are identified.

Heterogeneities in Face Recognition

Images used in face recognition are related to facial shape, skin, and hair. A 3D face image is related to the shape only. It is captured by a range-measuring system usually made from a laser range system or stereo vision system. Represent a range image taken from a viewpoint by $z(x, y)$. The pixel values measure the distances of the sensor to the facial surface points.

In developing face biometric engines and systems using spectral images, researchers and engineers have identified intrinsic and extrinsic factors that affect face recognition. The ► Lambertian law provides an image formation model, relating a spectral image with the 3D

shape of the sensed object, the object surface properties, and the illumination source:

$$I(x, y) = \rho(x, y)\mathbf{n}(x, y)\mathbf{s} \quad (1)$$

where $I(x, y)$ is the spectral image, $\rho(x, y)$ is the albedo of the facial surface material at point (x, y) (also a function of the illumination wavelength), $\mathbf{n} = (n_x, n_y, n_z)$ is the surface normal (a unit row vector) at the 3D surface point $z(x, y)$, and $\mathbf{s}=(s_x, s_y, s_z)$ is the point lighting direction (a column vector, with magnitude). The normal directions $\mathbf{n}(x, y)$ may be derived from the range image $z(x, y)$, but not vice versa.

In developing face biometric engines and systems using spectral images, researchers and engineers have identified intrinsic and extrinsic factors that affect face recognition. The facial albedo and the surface shape are intrinsic factors pertinent to the face identity. These should be the most important information to be used for face recognition. On the other hand, extrinsic factors include illumination, facial ware, hairstyle, expression, and posture. Since they are irrelevant to the identity of the face, their influence on face recognition should be mitigated. Much research and development effort has been spent to minimize the impact of extrinsic factors, but the problems still persist and are difficult to solve [1].

Heterogeneous spectral face images have different albedos, and hence, encode intrinsic factors in different ways even if extrinsic factors are not accounted for. Set aside extrinsic factors and focus on the intrinsic ones. Given a still, frontal face under a fixed illumination, heterogeneous image formation processes produce face images of different image configurations of pixel values. The pixel values have different properties and interrelationships across heterogeneous face images.

While the above heterogeneities in HFBs are considered in the true sense, HFBs in the broad sense deals with heterogeneities in homogeneous face images captured under heterogeneous conditions. The VIS type of face images, for example, can be captured

- under different illumination conditions,
- by different types of image sensors, such as CCD and CMOS, or sensor brands,
- in different image resolutions,
- in different image quality, and
- by photo scanning or face sketching,

These cause heterogeneities in image formation and pixel configuration. Among these, face sketches may have different image styles and contain more heterogeneities. Also, heterogeneities due to image resolutions and different image quality also count. Quality control by imposing constraints on image acquisition conditions is thus suggested, for example, in the ISO/ICAO (International Organization for Standardization/International Civil Aviation Organization) standard [10].

Research Issues

Despite the heterogeneities, it is desired to perform face biometric identification and verification with whatever types of face images available. Research problems in HFBs include the following:

- Understanding heterogeneous image formation models: This provides a physical basis for modeling properties of heterogeneous face images.
- Discovering relationships between heterogeneous images: Relations or correlations between heterogeneous images of faces or sets of features derived thereafter may be discovered using heterogeneous image formation models.
- Formulating transformation of one type to another: With latent correlations discovered, one could construct a transformation or mapping from one type to another.
- Extracting common features: Discovered latent correlations could also be used for extracting common features for characterizing face identities in heterogeneous images.
- Matching across heterogeneous images: Matching algorithms should be developed based on extracted features that associate heterogeneous face image properties.
- Fusion of heterogeneous information: HFBs can take advantage of heterogeneous information in face images and fuse them to improve the performance.

Statistical learning can be used to develop algorithms for solving these problems. For example, for the recovery of face shape from a single NIR face image [8], for matching between VIS face and face sketch [5], VIS face and NIR face [6], and between 3D face and NIR face [7].

HFBs require the extraction of features common across heterogeneous types of faces so as to create a common ground for things to be compared. The extraction of common features is the most distinct issue

among all in HFBs, while the other issues have been researched in homogeneous face biometrics, and their results can be applied herein. Methods differ in how and where to extract the common features – whether it is done at one end of the two types as in the ► analysis-by-synthesis approach, or somewhere in the middle as in the ► common feature approach.

For example, in [6, 7], common features are extracted using canonical correlation analysis (CCA) [?] and are in the middle of the two ends rather than at a single end. In [5], VIS face images (at one end) are synthesized from face sketches (the other end) explicitly, and the matching is performed using features extracted from the VIS end of face images. Possible solutions to HFBs can therefore be categorized under two classes: common feature based and analysis-by-synthesis based.

Summary

Heterogeneous face biometrics (HFBs) perform biometric matching across heterogeneous face images. This article has discussed an analysis of problems in HFBs, identified issues therein, and point out research directions. HFBs could be used as a standalone module for biometric authentication or work as an added module to improve face recognition with homogeneous face images. HFBs are not only new directions for face-based biometrics, but also address the underlying issues in conventional homogeneous face biometrics in the broad sense of HFBs. Research and development of HFBs, with investigation into problems caused by heterogeneities in homogeneous face biometrics, may lead to better solutions.

Related Entries

► 3D Based Face Recognition
► Face Recognition, Near-Infrared
► Face Recognition, Thermal
► Hyperspectral and Multispectral Biometrics
► Skin Spectroscopy

References

1. Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A.: "Face recognition: A literature survey". ACM Computing Surveys 399–458 (2003)
2. Li, S.Z., Chu, R., Liao, S., Zhang., L.: Illumination invariant face recognition using near-infrared images. IEEE Trans. Pattern Anal. Mach. Intell. **26** (Special issue on Biometrics: Progress and Directions), 627–639 (2007)
3. Kong, S.G., Heo, J., Abidi, B., Paik, J., Abidi, M.: Recent advances in visual and infrared face recognition - A review. Comput. Vis. Image Underst. **97**(1), 103–135 (2005)
4. Bowyer, K.W., Chang, Flynn, P.J.: A survey of 3D and multi-modal 3d+2d face recognition. In: Proceedings of International Conference on Pattern Recognition, 358–361 (2004)
5. Tang, X., Wang, X.: Face sketch recognition. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 50–57 (2004)
6. Yi, D., Liu, R., Chu, R., Lei, Z., Li, S.Z.: Face matching between near infrared and visible light images. In: Proceedings of IAPR International Conference on Biometric, Seoul, Korea (2007)
7. Yang, W., Yi, D., Lei, Z., Sang, J., Li, S.Z.: 2D-3D face matching using cca. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, Amsterdam, The Netherlands (2008)
8. Lei, Z., Bai, Q., He, R., Li, S.Z.: Face shape recovery from a single image using cca mapping between tensor spaces. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2008)
9. NIST: Multiple Biometric Grand Challenge (MBGC). http://face.nist.gov/mbgc (2008)
10. ANSI/INCITS/ISO/IEC 19794-5: Information Technology - Biometric Data Interchange Formats - Part 5: Face Image Data [S] (2007)

Heterogenous Face Image Matching

► Heterogeneous Face Biometrics

Hidden Markov Models

Javier Hernando
Technical University of Catalonia, Barcelona, Spain

Synonym

HMM

Definition

A Hidden Markov Model is a twofold stochastic process composed by a first-order Markov chain, which is

a finite state machine ruled by state transition probabilities that solely depend on the immediate predecessor, and an associated probabilistic function. In a regular Markov model, the state is visible to any observer external to the model, whereas in a *Hidden* Markov Model (HMM), the state is not observable, that is, hidden, but each state has associated output probabilities over the possible observable tokens.

Introduction

Hidden Markov Models (HMMs) [1] were introduced by L.E. Baum in the late 1960s [2], and since the mid-1970s [3–5], they have become popular to model the statistical variation of the spectral features in speech recognition research. In the late 1980s, HMMs were applied to the analysis of DNA and other biological sequences. Nowadays, they are ubiquitous in bioinformatics, and in particular, in biometrics.

Architecture and Types

An HMM is characterized by the following components:

- A set of N states of a first-order Markov chain $S = \{S_i\}$, $i = 1, \ldots, N$. Denoting the instants of time regularly spaced associated with the state transitions by $t = 1, 2, \ldots, T$, the state in time t is denoted by q_t.
- The set of transition probabilities between the states. Assuming the first-order Markov chain condition, they can be represented by the matrix $A = \{a_{ij}\}$, $i, j = 1, \ldots, N$, where

$$a_{ij} = P(q_{t+1} = S_j / q_t = S_i) \quad 1 \leq i, j \leq N$$

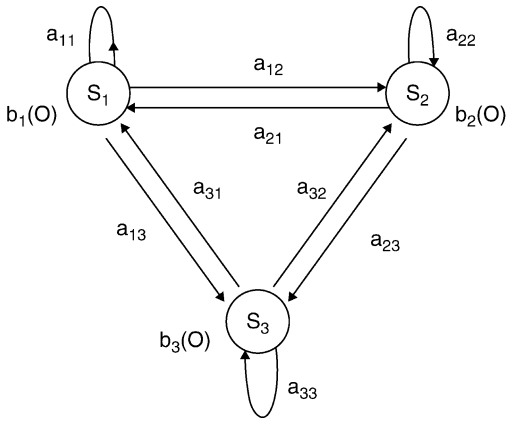
- The initial state probability matrix, which will be denoted by the vector

$$\pi_i = P(q_1 = S_i) \; 1 \leq i \leq N$$

- The output probabilities $B = \{b_j(O_t)\}$, $j = 1, \ldots, N$, where

$$b_j(O_t) = P(O_t | q_t = S_j) \; j = 1, \ldots, N,$$

b_j being the probability distribution corresponding to the state j, assumed to be independent of time, and O_t the value of the observation at instant t, corresponding to the observation sequence $O = \{O_t\}$, $t = 1, \ldots, T$.



Hidden Markov Models. **Figure 1** Representation of a Hidden Markov Model.

Therefore, the HMM can be represented as $\lambda = (A, B, \pi)$. The architecture of a three state HMM is illustrated in Fig. 1. When any state is reachable from any other state, as it is represented in the figure, the model is called ergodic. However, in text-dependent speaker recognition, as well as in speech recognition, the models are left-to-right, that is, the states are only reachable from the state itself of lower index states.

The probability distributions of the output probabilities are discrete in the so-called Discrete Hidden Markov Models (DHMM) and continuous in the so-called Continuous Hidden Markov Models (CHMM). When the original observation data are continuous, they must be quantized if the DHMM is preferred. In the case of speech and speaker recognition, where the observation data are a sequence of acoustic parameter vectors, vector quantization techniques were used when the computational efficiency and amount were crucial. However, nowadays CHMMs have become most popular in speech and speaker recognition.

In the DHMMs, the output probabilities take values in a finite set of symbols called alphabet $V = \{v_k\}$, $k = 1, \ldots, M$, M being the alphabet size. Hence, they can be denoted with the matrix $B = \{b_j(k)\}$, where

$$b_j(k) = P(v_k \text{en } t | q_t = S_j), \quad j = 1, \ldots, N, \; k = 1, \ldots, M.$$

In the CHMMs, generally, the output probabilities take values in multidimensional, continuous space, and they are modeled by parametric multivariate probability density functions. A Gaussian mixture is the most used [6, 7], because it can approximate any probability density function with an adequate number of mixtures. In this case,

$$b_j(O_t) = \sum_{m=1}^{M} c_{jm} \aleph(O_t, \mu_{jm}, \Sigma_{jm}),\ j = 1, \ldots, N$$

that is, a linear combination with weight c_{jm} of M ℵ multivariate Gaussian probability density functions with mean vector μ_{jm} and covariance matrix Σ_{jm}.

In both cases, once the number of states and the permitted transitions between them (topology) and the parameter tying between states are predefined, these three main problems must be solved in order to use HMMs in real applications.

Evaluation

Given the observation sequence $O = \{O_t\}$, $t = 1, \ldots, T$ and a model $\lambda = (A, B, \pi)$, the evaluation problem consists in efficiently computing $P(O|\lambda)$, the probability of the observation sequence, given the model. This probability can be used to classify observation sequences in recognition applications.

The brute force solution for this problem is to enumerate all possible state sequences and calculate their score directly. For a fixed state sequence $Q = \{q_t\}$, $t = 1, \ldots, T$ and assuming independence between the observations O_t, the probability of the observation sequence O can be written as

$$P(O|O, \lambda) = \prod_{t=1}^{T} P(O_t|q_t, \lambda) = b_{q1}(O_1) b_{q2}(O_2) \cdots b_{qT}(OT).$$

On the other hand, the probability of the state sequence Q can be written as

$$P(Q|\lambda) = a_{q1} a_{q1q2} a_{q2q3} \cdots a_{qT-1qT}.$$

Finally, the joint probability of both the observation and a fixed state sequences given the model $P(O, Q|\lambda)$ is simply the product of the terms shown earlier, and the probability of the observation sequence given the model can be obtained summing this product over all possible state sequences

$$P(O|\lambda) = \sum_{q1q2\cdots qT} \pi_{q1} b_{q1}(O_1) a_{q1q2} b_{q2}(O_2) \cdots a_{qT-1qT} b_{qT}(O_T).$$

This procedure has a time complexity of $o(2TN^T)$, and therefore it is not tractable in real application, even for moderate values of N and T. Fortunately, there is an alternative recursive procedure called *Forward–Backward* [2] that computes this probability in an efficient way, which is summarized as follows.

The forward variable be α_1(i) is given by

$$\alpha_t(i) = P(O_1 O_2 \cdots O_t, q_t = S_i|\lambda).$$

It is easy to show that it can be computed in the following inductive way:

1. Initialization $\alpha_1(i) = \pi_i b_i(O_1) i = 1, \ldots, N$
2. Induction $\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), t = 1, \ldots T-1, j = 1, \ldots, N$
3. Termination $P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$

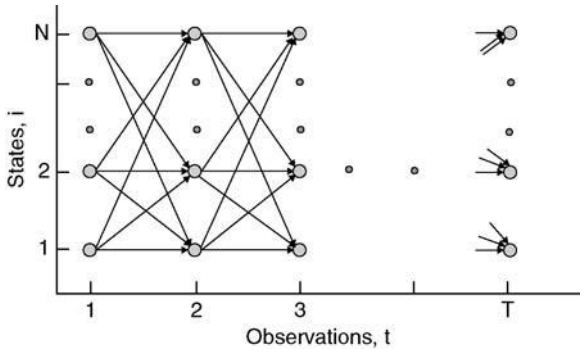
Theses computations can be organized in the observations and state lattice as shown in Fig. 2, and have a time complexity of $o(N^2T)$, which is acceptable for real application.

Alternatively, the backward variable

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, \ldots, O_T|q_t = S_i, \lambda).$$

It can be computed in the following inductive way:

1. Initialization $\beta_T(i) = 1,\ i = 1, \ldots, N_N$
2. Induction (see ure 3) $\beta_t(i) = \sum_{j=1} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), t = T-1, \quad T-2, \ldots, 1 \quad i = 1, \ldots, N$
3. Termination $P(O|\lambda) = \sum_{i=1}^{N} \pi_i b_i(O_1) \beta_1(i)$



Hidden Markov Models. **Figure 2** Computation lattice of the Forward algorithm.

Decoding

Given the observation sequence $O = \{O_t\}$, $t = 1, \ldots, T$ and a model $\lambda = (A, B, \pi)$, the decoding problem consists in choosing the *opitmal* state sequence $Q = \{q_t\}$, $t = 1, \ldots. T$, which best *explains* the observations. The solution of this problem permits to obtain information about the hidden process and is also a good and efficient approximation of the evaluation problem.

The most popular criterion is to select the state sequence that maximizes the conditional probability

$$Q^* = \arg\max_Q \{P(Q|O, \lambda\} = \arg\max_Q \{P(Q, O|\lambda)\}.$$

Again, the brute force method cannot be tackled, even for moderate values of N and T, and in this case, it is replaced by the ▶ Viterbi algorithm, which recursively solves the problem in an efficient way.

The Viterbi algorithm defines the variables

$$\delta_t(i) = \max_{q1,q2,\ldots,qt} \{P(q_1 q_2 \cdots q_t = i, O_1\ O_2 \cdots O_t|\lambda)\}$$

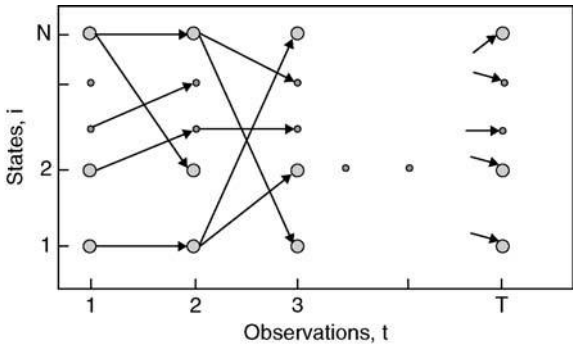
and $\psi_t(j)$ in order to recursively maximize the conditional probability and to retrieve the corresponding optimal state sequence respectively. Both of them are updated by using the following procedure:

1. Initialization: $\delta_t(i) = \pi_i\, b_i(O_1)\ i = 1,\ldots, N\, \psi_t(j) = 0$
2. Recursion: $\delta_t(j) = \max_{i=1,\ldots,N} \{\delta_{t-1}(i)a_{ji})\}b_j(O_t)$, $t = 2, \ldots, T, j = 1, \ldots, N \psi t(j) = \arg\max_{i=1\ldots N} \{\delta_{t-1}(i)a_{ij})\}\ t = 2, \ldots, T,\ j = 1, \ldots, N$
3. Termination: $P^* = \max_{i=1\ldots N} \{\delta_{t-1}(i)a_{ji})\}$
4. State sequence backtracking: $qt^* = \arg\max_{i=1\ldots N} (\delta_T(i)\}\ q_t^* = \psi_{t+1}(q_{t+1}^*)\ t = T-1,\ T-2, \ldots, 1$

In this case also, the lattice structure implements the computations in an efficient way (see Fig. 3). It is worth noting that not all the state transitions are considered, but only the ones that lead to the maximum probability.

Estimation

Given the observation sequence $O = \{O_t\}$, $t = 1, \ldots, T$, the estimation problem consists in adjusting the parameters of the model $\lambda = (A, B, \pi)$ so as to maximize the probability of observation of this sequence, given the model $P(O|\lambda)$. The solution to this problem permits to develop a method to train self-learning classifiers.



Hidden Markov Models. Figure 3 Computation lattice of the Viterbi algorithm.

This is the most challenging of the three problems. In fact, given a finite sequence of observations, it is not possible to optimally estimate the model parameters. However, the model parameters can be chosen to locally maximize the probability $P(O|\lambda)$ by using the ▶ Baum–Welch algorithm [8], which is summarized in the following section.

For the DHMM case, the variable

$$\xi_t(i,j) = P(q_t = S_i, q_{t+1} = S_j|O, \lambda) = \frac{P(q_t = S_i, q_{t+1} = S_j, O|\lambda)}{P(O|\lambda)},$$

which is the probability of being in the state S_i at time t and in state S_j at time t + 1, given the model and the observation sequence. This value can be written in terms of the forward and backward variables an the parameter models as

$$\xi_t(i,j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i,j=1}^{N} \alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}$$

The variable

$$\lambda_t(i) = P(q_t = S_i|O, \lambda) = \frac{P(q_t = S_i, O|\lambda)}{P(O|\lambda)}$$

which is the probability of being in the in the state S_i at time t, given the model and the observation sequence, and can also be written as

$$\lambda_t(i) = \frac{\alpha_t(\mathrm{i})\beta_t(\mathrm{i})}{\sum_{i=1}^{N} \alpha_t(\mathrm{i})\beta_t(\mathrm{i})}$$

The sum of $\xi_t(i, j)$ over $t = 1, \ldots, T-1$ can be interpreted as the expected value (over time) of the number of transitions from the state S_i to the state S_j. Furthermore, the sum of $\gamma_t(i)$ over $t = 1, \ldots, T$ is the expected number of times that state S_i is visited, the sum of $\gamma_t(i)$ over $t = 1, \ldots, T-1$ is the expected number of transitions made from the state S_i, and the sum of $\gamma_t(i)$ over $t = 1, \ldots, T$, when the symbol is v_k observed, is the expected number of times that the model generates the symbol v_k in the state S_i. Therefore, reasonable re-estimation formulas for π, A and B are

$$\bar{\pi}_i = \lambda_1(i) \quad i = 1, \ldots, N$$

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(\mathrm{i}, \mathrm{j})}{\sum_{t=1}^{T-1} \gamma_t(\mathrm{i})}, \quad i, j = 1, \ldots, N$$

$$\bar{b}_j(k) = \frac{\sum_{t=1, Ot=vk}^{T} \gamma_t(\mathrm{j})}{\sum_{t=1}^{T} \gamma_t(\mathrm{j})}, \quad j = 1, \ldots N, \quad k = 1, \ldots M$$

For the CHMMS, the variable $\gamma_t(j, k)$ is given by

$$\lambda_t(j, k) = \frac{\alpha_t(j)\beta_t(j)}{\sum_{j=1}^{N} \alpha_t(j)\beta_t(j)} \frac{c_{jk}\aleph\left(O_t, \mu_{jk}, \Sigma_{jk}\right)}{\sum_{m=1}^{M} c_{jm}\aleph\left(O_t, \mu_{jm}, \Sigma_{jm}\right)}$$

which is the probability of being in the state j at time t with the kth mixture accounting for O_t. Following the analogous reasoning done for DHMMS, a reasonable re-estimation formula for B is (formulas for π and A are the same)

$$\bar{c}_{jk} = \frac{\sum_{t=1}^{T} \gamma_t(\mathrm{j}, \mathrm{k})}{\sum_{t=1}^{T} \sum_{k=1}^{M} \gamma_t(\mathrm{j}, \mathrm{k})}, \quad j = 1, \ldots N, \quad k = 1, \ldots M$$

$$\bar{\mu}_{jk} = \frac{\sum_{t=1}^{T} \gamma_t(\mathrm{j}, \mathrm{k}) O_t}{\sum_{t=1}^{T} \sum_{k=1}^{M} \gamma_t(\mathrm{j}, \mathrm{k})}, \quad j = 1, \ldots, N, \quad k = 1, \ldots, M$$

$$\bar{\Sigma}_{jk} = \frac{\sum_{t=1}^{T} \gamma_t(\mathrm{j}, \mathrm{k}) \left(O_t - \mu_{jk}\right) \left(O_t - \mu_{jk}\right)'}{\sum_{t=1}^{T} \sum_{k=1}^{M} \gamma_t(\mathrm{j}, \mathrm{k})},$$

$$j = 1, \ldots, N, \quad k = 1, \ldots, M$$

where prime denotes vector transpose.

It can be shown that using this re-estimation formulae uniformly converge to Maximum Likelihood estimation, but they only lead to local maxima and that, in most applications, the optimization surface is complex, and hence initialization is an issue.

On the other hand, these formulae can also be obtained directly by maximizing the Baum auxiliary function

$$Q(\lambda, \lambda') = \sum_{Q} P(Q|O, \lambda) \log[P(Q|O, \lambda')]$$

and also they can be interpreted as an implementation of the Expectation-Modification algorithm [9].

Finally, it is worth noting that this optimization problem has been solved by conventional gradient techniques. This approach permits the use of other optimization criteria like MMI (Maximal Mutual Information) [10]. Furthermore, less formalized algorithms like corrective training [11] have been applied.

Summary

Hidden Markov Models composed by a hidden Markov chain and an observable associated probabilistic function have become ubiquitous in biometric in the recent years. Once the topology has been designed, the three conventional procedures, which solve the main problems for the application of the Hidden Markov Models in the real world, have been discussed in this paper: the ▶ Forward-Backward algorithm for the evaluation of the model, the Viterbi algorithm for the decoding of the optimal sequence of the states, and the Baum–Welch algorithm to adjust the model parameters to the observations.

Related Entries

▶ Biometric Algorithm
▶ Classifier Design
▶ Gaussian Mixture Models

- ▶ Machine-Learning
- ▶ Probability Distribution

References

1. Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recognition. Proc. IEEE **77**, 257–286 (1989)
2. Baum, L.E., Egon, J.A.: An inequality with applications to statistical estimation for probabilistic functions of a markov process and to a model for ecology. Bull. Am. Meteorol. Soc. **73**, 360–363 (1967)
3. Baker, J.K.: The DRAGON system – an overview. IEEE Acoust. Speech Signal Process. Mag. **23**, 24–29 (1975)
4. Bakis, R.: Continuous speech recognition via centisecond acoustic states. In: Proceedings of the ASA Meeting (1976)
5. Jelinek, F.: Continuous speech recognition by statistical methods. Proc. IEEE **64**, 532–556 (1976)
6. Rabiner, L.R., Juang, B.H.: An introduction to hidden markov models. IEEE Acoust. Speech Signal Process. Mag. **3**, 4–16 (1986)
7. Juang, B.H., Levinson, S.E., Sondhi, M.M.: Maximum likelihood estimations for multivariate mixture observations of markov chains. IEEE Trans. Inf. Technol. **32**, 307–309 (1986)
8. Baum, L.E.: An inequality and associated maximization technique in statistical estimation of probabilistic functions of markov processes. Inequal. **3**, 1–8 (1972)
9. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. **39**, 1–38 (1977)
10. Merialdo, B.: Phonetic recognition using hidden markov models and maximum mutual information training. In: Proceeding of the International Conference on Acoustics, Speech and Signal Processing pp. 111–114 (1988)
11. Bahl, Brown, P.F., de Souza, P.V., Mercer, R.L.: Estimating hidden markov model parameters so as to maximize speech recognition accuracy. IEEE Tran Speech Audio Processing **1**, 77–83 (1993)

Hill-Climbing Attack

Security attacks based on generating artificial data, injecting it in the system and after analyzing the output, modify such data, as to improve the output. This is done recursively till the output is the desired result. In biometrics this attack can be used to generate a synthetic sample, by analyzing the matching score returned by the system.

- ▶ Tamper-proof Operating System

Histogram Equalization

Histogram equalization is a common technique for adjusting the pixel intensities of images. In histogram equalization, a monotonic transformation is applied to the intensities of the pixels in an image. The transformations are chosen so that the resulting images have a standard or specified histogram. This is sometimes performed to enhance the contrast or reduce the effect of lighting variation on the appearance of a scene.

- ▶ Face Variation
- ▶ Image Pattern Recognition
- ▶ Illumination Compensation
- ▶ Pre-Processing

HMM

- ▶ Hidden Markov Models

Human Computing

Human computing is the merging of mobile communications and sensing technologies, with the aim of enabling a pervasive and unobtrusive intelligence in the surrounding environment supporting the activities and interactions of the users in a human-centered manner. Specifically, human-centered interfaces support detection of subtleties of and changes in the user's behavior, and initiate interactions based on this information rather than simply responding to the user's commands. Technologies like machine analysis of facial expressions and affective computing are inherent human-computing technologies.

- ▶ Facial Expression Recognition

Human Dental Atlas

Human dental atlas is a model for describing shape and relative positions of teeth in a fully developed adult dentition. The model categorizes a complete set of 32 teeth into 6 different classes according to tooth shape and position.

► Dental Biometrics

Human Detection and Tracking

James W. Davis, Vinay Sharma, Ambrish Tyagi, Mark Keck
Ohio State University, Columbus, OH, USA

Synonyms

Association; Correspondence; Localization; Pedestrian detection; Target detection; Video surveillance

Definition

Human detection and tracking are tasks of computer vision systems for locating and following people in video imagery. Human detection is the task of locating all instances of human beings present in an image, and it has been most widely accomplished by searching all locations in the image, at all possible scales, and comparing a small area at each location with known templates or patterns of people. Human tracking is the process of temporally associating the human detections within a video sequence to generate persistent paths, or trajectories, of the people. Human detection and tracking are generally considered the first two processes in a video surveillance pipeline, and can feed into higher-level reasoning modules such as action recognition and dynamic scene analysis.

Introduction

In relation to large-scale biometric and video surveillance systems, there is an increasing need for persistent, autonomous sensing of people in video using computer vision algorithms. Biometric systems can be employed in close proximity (using fingerprint or iris scans) or remotely (employing face or gait patterns) to identify or confirm a person's identity. Hence knowing "when" and "where" to attempt a remote biometric signature can greatly aid the task in terms of reliability and efficiency. Being able to detect and track people from a distance using video cameras can be used to reliably cue remote biometric sensors to engage only when appropriate. Furthermore, the desire for (semi-) autonomous video surveillance necessitates computer vision systems to detect, track, and analyze the behaviors of people in video. For example, a large airport security system could use computer vision systems to simultaneously monitor multiple video camera feeds and could present only those camera feeds showing suspicious activity to security personnel.

Fundamental to the aforementioned remote biometrics and video surveillance applications is the ability to automatically detect and track people in video. The purpose of human detection is to find the location of every person in each video image, while producing as few false detections as possible. The most common methods used for detection are template or pattern matching algorithms. The detections themselves can be used for a variety of applications, such as person counting or determining if a person is located in an area where nobody should be present. Human tracking is the process which associates the human detections over time to create a consistent path trajectory for each person. Tracking systems typically retain some history of the past detections and use this information along with the current detection to match and update the trajectories. The output of the human detection and tracking systems can be further used in higher-level reasoning modules such as activity analysis and recognition.

Human Detection

Detection, in general, refers to the task of determining whether or not an instance of a specific object class is present in the scene. With the growing emphasis on biometrics and surveillance, special attention has been paid to the task of detecting humans in images. Smart urban surveillance systems are typically designed to monitor people, both to ensure safety and to recognize unlawful acts. The success of such surveillance systems

is critically dependent on their ability to first reliably detect and localize all humans in the scene. The task of human detection is especially challenging due to the non-rigid, articulate nature of the human body.

The typical detection output is a bounding box surrounding each person in the image (Fig. 1). Most researchers approach human detection as a pattern classification problem, where the emphasis is on learning characteristic appearance features of humans from specially designed training datasets. These datasets generally consist of hundreds of labeled images containing people (positive examples) and a larger collection of images not containing people (negative examples). Often the positive examples included in the datasets are constrained to specific poses or camera views such that the variability of the human class is limited to some extent. Within such a training framework, several researchers in human detection focus on devising the most appropriate set of features that would encompass the variations exhibited naturally by humans (e.g., changes in pose, size, appearance, non-rigid motion, etc.) and at the same time adequately differentiate the human class from the rest of the objects in the scene. Image features such as Haar wavelets [1], histograms of oriented gradients [2], and covariance matrices [3], have shown to provide promising results.

Research in human detection also focuses on efficient learning algorithms and classification techniques that provide high detection accuracy with low false positive rates. For example, the use of AdaBoost to effectively combine hundreds of weak, computationally inexpensive classifiers was introduced for human detection in [4]. In [3] a classification scheme was proposed that takes advantage of the geometry of the manifold on which the extracted features reside. Other approaches adopt statistical tools such as field models for modeling human shape and motion [5] and support vector machines for human classification [2]. Certain approaches to human detection also make explicit use of the structure and ▶ kinematics of the human body, utilizing some form of human body model [6]. Such techniques often detect the human body by its smaller components (arms, legs, torso) instead of training a classifier for the whole-body form.

Human Detection and Tracking. **Figure 1** Typical human detection output showing the location and scale of people.

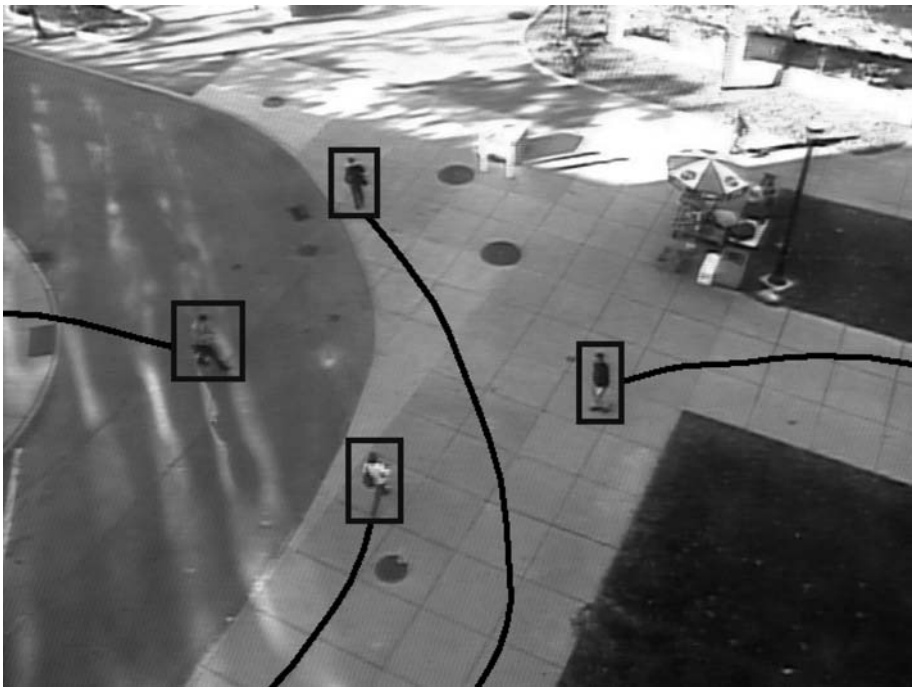
Human detection algorithms can also be differentiated based on whether or not they utilize motion cues. Humans have movement patterns that are distinct from other object classes, and some algorithms exploit motion features such as periodicity and motion symmetry to detect humans in video [7]. While motion provides a very powerful cue to detect humans, it is important for human detection algorithms to utilize both appearance and motion features in a balanced manner, so as to detect people irrespective of whether they are moving or stationary [2, 4, 8].

Typically human detection algorithms aim at providing only the location and scale of the people in the scene. More recently there has been a growing emphasis on additionally recovering the silhouette shape of each detected person. Such algorithms typically require training images in which the silhouettes are manually annotated in each image [9], though there are other algorithms that require either only very limited or no manual annotation of the training dataset [8]. Since humans have articulated limbs, can occur in different poses, and may carry objects (e.g., briefcase, backpack, etc.), acquiring the actual silhouette shape can provide valuable cues for gait and activity recognition systems.

Human Tracking

Tracking is the process of temporally associating the location of the target object (e.g., human) from one frame to the next in a sequence of images (video). A detailed survey on object tracking can be found in [10]. Human tracking is of considerable interest to security applications, and the tracking information is useful for determining the origin, pathway, and destination of each person (Fig. 2). As most human tracking approaches employ standard (generic) object tracking algorithms, an overview of these techniques is presented here.

Most tracking algorithms can be broadly classified into two categories: (1) data association and filtering techniques or (2) target representation and localization methods. The first category is a class of solutions characterized by associating a set of detection locations from the previous image to a set of detections in the current image. The second category is distinguished by building a target model of the appearance of the object in the initial frame and searching for the object location in successive frames by generating candidate models and finding the best possible target-candidate match. The tracking task is complicated by



Human Detection and Tracking. **Figure 2** Trajectories of tracked people.

changes in illumination, image noise, nonrigid target motion, and occlusion. For practical applications, algorithms from both categories may be combined to obtain robust and efficient trackers.

Data Association and Filtering

Tracking techniques based on association assume that there is a set of detection locations in an image at time t and a set at time $t+1$. The goal of these trackers is to find a mapping from the set at time t to the subsequent set at time $t+1$, establishing detection correspondence between the frames. The simplest solution is to map the set at time t to its successor according to Euclidean distance. A more complex approach may incorporate the velocity of each detection into the computation, assuming both positional proximity and smooth continuity in the movement [11]. These data association techniques use simple assumptions and minimize some criteria according to kinematic features extracted from the video sequence. Naturally occlusion is a significant challenge with these techniques, as the mapping is more difficult to recover when the cardinalities of the two sets are not equal (i.e., missing data).

Filtering strategies improve upon the basic data association approach by supplying a specific model of object motion and algorithms by which one can recover the optimal sequence of locations of the object throughout the video. These strategies generally follow a predict and update framework, where the new position and velocity of the object at time $t+1$ is hypothesized/predicted by the model from time t, and then the model is updated after the frame at time $t+1$ is processed. These methods tend to smooth the trajectories (removing noise) and improve the robustness to occlusion (as new object locations are hypothesized even when data are not present).

One of the simplest and most common of these filtering approaches is the Kalman filter [12]. The Kalman filter is a state-space model that makes assumptions that the position and velocity at time $t+1$ is a linear function of the position and velocity of the object at time t plus some additive Gaussian noise and that the posterior distribution of the hidden state variable is Gaussian. These assumptions allow the model to recover the optimal positions and velocities from the data locations (assuming the object motion obeys the model assumptions). Unfortunately trajectories of human motion can at times be nonlinear, and in such cases the Kalman filter framework is likely to fail. Extended Kalman filters have been proposed to relieve the linear model assumption by taking the first-order approximation of a non-linear process and casting the problem in a similar predict-update framework [13].

Particle filtering is a general filtering strategy that relieves the previous Gaussian assumption, allowing the posterior distribution to be generic. This is achieved by representing the posterior distribution as a set of weighted points (particles) in the state space. This representation is predicted at time $t+1$ by sampling an importance density function (which approximates the posterior) at time t. It is then updated at each frame by changing the weight of each sample, thus allowing the modes of the density to propagate by assimilating new observations. This type of filtering was popularized by [14], where particle filters were employed to track objects through cluttered scenes.

The drawback to most of these algorithms is that in most cases the models do not incorporate appearance information into the system, relying strictly on kinematic data (position and velocity of the detections). However, there are many examples where a tracker that could take advantage of appearance information would significantly increase the tracking performance, for instance when tracking a person wearing a red shirt walking in front of a white wall. The next category of trackers addresses this type of tracking.

Target Representation and Localization

Target representation and localization tracking frameworks initially extract an appearance model of the target object from the first image, and then search a series of candidate locations in the next image to find the best matching candidate. Unlike the data association and filtering methods, this category of tracking algorithms is mostly a bottom-up process, and essentially performs detection in each frame. The algorithms must be able to handle appearance changes of the target over time as it is being tracked.

Popular features used for target representation include color histograms, image gradients, and covariance matrices (of features) within an image patch containing the target object. Localization techniques varying

from global exhaustive search to local heuristic optimizations have been employed to search for the target locations in subsequent image frames. Many localization algorithms benefit from exploiting a spatio-temporal locality constraint, which assumes that the location of the object from one frame to the next changes gradually. This constraint helps reduce the target search space and can thus result in faster, real-time tracking algorithms.

Mean shift, a nonparametric density gradient estimator [15], is a common method used to track objects by finding the mode (peak) of the similarity surface generated by comparing the object appearance model (e.g., color histogram) with the target candidates. Similarity is evaluated as the Bhattacharyya coefficient between the model and candidate distributions. This algorithm performs a local optimization on the search surface starting from the previously known object location and is well known for its computational efficiency (real-time frame rates can be obtained). Robust tracking results are obtained under variable environmental (illumination, occlusions), object (articulate, nonlinear motion), and camera (static, moving, jitter) configurations.

The use of covariance features for the target representation was proposed by [3]. The covariance matrix of features (e.g., position, intensity, color, gradients) extracted from an image patch enables a compact representation of both the spatial and the statistical properties of the object. The tracker performs a search in the image by comparing the given covariance model with the covariance matrix at each possible location using an appropriately defined distance metric. The location which is most similar to the target model is assigned to be the new target position in the image.

Summary

Human detection algorithms detect the presence of people in imagery and must accommodate all of the appearance variations while not selecting non-human entities. Human tracking temporally associates the human detections within video sequences to generate trajectories, and is complicated by short- and long-term occlusions. By providing information regarding the location and movement of humans in the scene, human detection and tracking algorithms enable applications such as remote biometrics and video surveillance.

Related Entries

▶ Image Pattern Recognition

References

1. Oren, M., Papageorgiou, C., Sinha, P., Osuma, E., Poggio, T.: Pedestrian detection using wavelet templates. In: Proceedings of Computer Vision and Pattern Recognition (1997)
2. Dalal, N., Triggs, B., Schmid, C.: Human detection using oriented histograms of flow and appearance. In: Proceedings of European Conference on Computer Vision (2006)
3. Tuzel, O., Porikli, F., Meer, P.: Human detection via classification on riemannian manifolds. In: Proceedings of Computer Vision and Pattern Recognition (2007)
4. Viola, P., Jones, M., Snow, D.: Detecting pedestrians using patterns of motion and appearance. In: Proceedings of International Conference Computer Vision (2003)
5. Wu, Y., Yu, T.: A field model for human detection and tracking. IEEE Trans. Patt. Analy. and Mach. Intell. **28**(5), 753–765 (2006)
6. Ramanan, D., Forsyth, D., Zisserman, A.: Strike a pose: Tracking people by finding stylized poses. In: Proceedings of Computer Vision and Pattern Recognition (2005)
7. Lee, S., Liu, Y., Collins, R.: Shape variation-based frieze pattern for robust gait recognition. In: Proceedings of Computer Vision and Pattern Recognition (2007)
8. Sharma, V., Davis, J.: Integrating appearance and motion cues for simultaneous detection and segmentation of pedestrians. In: Proceedings of International Conference Computer Vision (2007)
9. Leibe, B., Seemann, E., Schiele, B.: Pedestrian detection in crowded scenes. In: Proceedings of Computer Vision and Pattern Recognition (2005)
10. Yilmaz, A., Javed, O., Shah, M.: Object tracking: A survey. ACM Comput. Surv. **38**(4) (2006)
11. Rangarajan, K., Shah, M.: Establishing motion correspondence. Comp. Vis. Graph. Img. Proc. **54**(1), 56–73 (1991)
12. Kalman, R.: A new approach to linear filtering and prediction problems. Trans. ASME-J. Basic Eng. **82**, 35–45 (1960)
13. Julier, S., Uhlmann, J.: A new extension to the kalman filter to nonlinear systems. In: SPIE AeroSense Symposium (1997)
14. Isard, M., Blake, A.: Condensation – conditional density propagation for visual tracking. Int. J. Comp. Vis. **29**(1), 5–28 (1998)
15. Comaniciu, D., Ramesh, V., Meer, P.: Kernel-based object tracking. IEEE Trans. Patt. Analy. and Mach. Intell. **25**(5), 564–577 (2003)

Human Factors

▶ Ergonomic Design for Biometric Systems

Human Movement, Psychology

Synonyms

Action Categorization; Action Understanding

Definition

The psychology of human movement is a broad ranging field that includes both how the motor control system produces movements, and how the sensory system perceives these movements itself and from others. Since both the structure of the body and the strategy for producing movements are unique they provide constraints that are potentially important for the sensory interpretation of movement. Applied areas of study in the psychology of human movement include sports psychology and social psychology, particularly when it applies to nonverbal communication interpretation of visual information from movements such as gait is of particular interest for biometrics. In the domain of visual perception, the psychology of human movement perception is becoming an increasingly important example of how the visual system processes a complex signal changing over time and attaches meaning and social significance to this signal.

► Psychology of Gait and Action Recognition

Human-Biometric Sensor Interaction (HBSI)

► Ergonomic Design for Biometric Systems

Human–Computer Interaction (HCI) and User Interfaces

Human–Computer interaction is the command and information flow that streams between the user and the computer. It is usually characterized in terms of speed, reliability, consistency, portability, naturalness, and users' subjective satisfaction. Human–computer interface (or simply "user interface") is a software application, a system that realizes human–computer interaction.

► Biometric System Ergonomic Design
► Facial Expression Recognition

Human-Interpretable Fingerprint Classes

Fingerprints are grouped based on some visual characteristics of fingerprint images determined by human experts. Such groups are called human-interpretable fingerprint classes. An excellent example of human-interpretable fingerprint classes is the well-known Galton–Henry classification scheme proposed by Sir Francis Galton and Edward Henry. The five most common Galton–Henry classes are called arch, tented arch, left loop, right loop, and whorl, which are easy to be understood even by ordinary people.

► Fingerprint Classification

Hypothesis Test

Hypothesis testing refers to the process of using statistical analysis to determine if the observed differences between two or more samples are due to random chance (as stated in the background hypothesis) or true differences in the samples (as stated in the target hypothesis).

► Universal Background Models

ICP Algorithm

Iterative Closest Point (ICP) algorithm developed by Besl and Mckay, is a well-known method to align 3D shapes. However ICP requires that every point in one set have a corresponding point on the other set. This cannot be guaranteed in practice. As a result modified ICP algorithms exist in the literature.

► Ear Biometrics, 3D

ID Photograph

► Photography for Face Image Data

Identification

Biometric identification is a process that ranks the biometric references in the enrolment database in order of decreasing similarity against a recognition biometric sample and then makes a decision, based on the similarity scores, about the identity w.r.t. to the references.

► Verification/Identification/Authentication/Recognition: The Terminology

Identity Level in the Speech Signal

Speech production is an extremely complex process, whose result depends on many variables at different levels, including sociolinguistic factors (e.g., level of education, linguistic context, and dialectal differences) and morphological issues (e.g., vocal tract length and shape or the dynamic configuration of the articulatory organs). These multiple influences will be simultaneously present in each speech act and some or all of them will contain specificities of the speaker. Hence, it is needed to clarify and clearly distinguish the different levels and sources of speaker information that should be extracted to model speaker individualities.

► Speaker Features

Identity Theft Reduction

► Fraud Reduction, Overview

Illumination

Ambient light sources may affect the appearance of a biometric image (such as face, fingerprint, or iris). The intensity and direction of these light sources can impact the performance of image-based biometric recognition algorithms.

► Biometrics, Overview
► Face Recognition, Near-Infrared
► Face Tracking
► Illumination Compensation
► Photography for Face Image Data

© 2009 Springer Science+Business Media, LLC

Illumination Compensation

Xudong Xie[1], Kin-Man Lam[2], Qionghai Dai[1]
[1]Automation Department, Tsinghua University, Beijing, China
[2]Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong, China

Synonyms

Lighting compensation; Illumination normalization

Definition

Due to difficulty in controlling the lighting conditions in practical applications, variable illumination is one of the most challenging tasks in face recognition. Prior to face recognition, illumination compensation has to be performed, whereby the uneven illumination of human faces is compensated and face images in normal lighting conditions are reconstructed. The reconstructed face images are then used for classification. An illumination compensation scheme includes the following modules: lighting category evaluation, shape normalization, and lighting compensation.

Introduction

Human face recognition, one of the most successful applications of image analysis and understanding, has received significant attention in the last decade. However, due to difficulty in controlling the lighting conditions in practical applications, variable illumination is one of the most daunting challenges in face recognition. As stated by Adini et al. [1], "The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity". Most of the available methods for face recognition, such as the *Principal Component Analysis* (PCA) [2], and the Independent Component Analysis (ICA) [3], encounter difficulties under varying lighting conditions. Some images under varying illuminations presented Fig. 1, all images are from the YaleB database [4]. It can be seen that although all these images are of the same person, due to the effect of uneven lighting, they look quite different. Therefore, when the images are under varying illumination, illumination compensation should be performed before face recognition. Here, "illumination compensation" means compensation for the uneven illuminations on human faces and reconstruction of face images in normal lighting conditions, then the reconstructed face images are used for classification.

Some algorithms have been proposed to reduce the effect of uneven lighting from an image processing point of view. Histogram equalization (HE) is a commonly used method to convert an image so that it has a uniform histogram, which is considered to produce an "optimal" overall contrast in the image. However, after being processed by HE, the lighting condition of an image under uneven illumination may sometimes become even more uneven. Adaptive histogram equalization (AHE) [5] computes the histogram of a local image region centered at a given pixel to determine the mapped value for that pixel leading to local contrast enhancement. However, the enhancement often leads to noise amplification in "flat" regions, and "ring" artifacts at strong edges. In addition, this technique is computationally intensive. Zhu et al. [6] proposed an illumination correction method, which uses an affine transformation lighting model based on a local estimation of the background and the illumination gain. However, the method is useful only when the images are under slowly varying illumination. The

Illumination Compensation. Figure 1 Samples of cropped faces under varying illuminations. The azimuth angles of the lighting of images from left to right column are: 0°, 0°, 20°, 35°, 70°, −50° and −70°, respectively. The corresponding elevation angles are: 20°, 90°, −40°, 65°, −35°, −40° and 45°, respectively.

Block-based Histogram Equalization (BHE) method [7] divides an image into a number of small blocks, and histogram equalization is performed within each of the image blocks. The BHE is simple, and the computation required is for less than that for AHE. However, as in the case of the AHE, noises are also enhanced after being processed by the BHE. The main idea of Local Normalization (LN) [8] is to split the face region into a set of triangular facets, and then the intensity values within each facet are normalized to be of zero mean and unit variance. This method is very fast, but it is also sensitive to variations caused by local shape distortions, such as expression variations. On the basis of a model-based method, [7] the general procedure for illumination compensation can be understood.

Human Face Model and Lighting Model

A face image is assumed to be a ► *Lambertian surface*, which can be described by the product of the albedo and the cosine angle between the point light source and the surface normal as follows:

$$I(x, y) = \rho(x, y)\mathbf{n}(x, y) \cdot \mathbf{s}, \tag{1}$$

where $I(x, y)$ is the intensity value observed of the pixel at (x, y) in the image, $0 \leq \rho(x, y) \leq 1$ is the corresponding albedo, $\mathbf{n}(x, y)$ is the surface normal direction, $\mathbf{s}$ is the light source direction, and its magnitude is the light source intensity. Suppose $I(x, y)$ and $I'(x, y)$ represent the pixel intensity values at (x, y) of the image under normal lighting conditions and the image under a certain kind of illumination, and $\mathbf{s}$ and $\mathbf{s}'$ are the corresponding light source directions, then the corresponding illumination ratio image [9] can be given as follows:

$$\begin{aligned} R_i(x, y) &= I'(x, y)/I(x, y) \\ &= (\rho(x, y)\mathbf{n}(x, y) \cdot \mathbf{s}')/(\rho(x, y)\mathbf{n}(x, y) \cdot \mathbf{s}) \\ &= (\mathbf{n}(x, y) \cdot \mathbf{s}')/(\mathbf{n}(x, y) \cdot \mathbf{s}) \\ &= A(x, y), \end{aligned} \tag{2}$$

where $A(x, y)$ is determined by the surface normal direction $\mathbf{n}(x, y)$ and the kind of illumination concerned. From (2), the following can be obtained:

$$I'(x, y) = A(x, y) \cdot I(x, y). \tag{3}$$

If the effect of additive noise at each point (x, y) is considered, the illumination model in (3) can be extended to the following:

$$I'(x, y) = A(x, y) \cdot I(x, y) + B(x, y), \tag{4}$$

where $A(x, y)$ and $B(x, y)$ denote the multiplicative noise and the additive noise for the pixel (x, y), respectively.

From (2) and (4), it can be seen that $A(x, y)$ and $B(x, y)$ are only determined by the shape information of a human face and the lighting condition, which hints us that if the shapes of identities are normalized to be the same, the values of $A(x, y)$ and $B(x, y)$ are invariant for different persons under the same lighting conditions. Therefore, after performing shape normalization, the values of $A(x, y)$ and $B(x, y)$ can be computed point by point for a certain illumination. Then these values can be used to undertake illumination compensation.

According to the illumination categories used in the YaleB database [4], the lighting conditions are divided into 65 categories. Each of the categories has different azimuth angles and elevation angles of the lighting. The azimuth angles in the database vary from $-130°$ to $+130°$, and the elevation angle ranges from $-40°$ to $+90°$. If both the azimuth angle and the elevation angle are equal to $0°$, it can be said that the subject is under normal illumination.

Shape Normalization and Lighting Compensation

Suppose that the pixel-wise correspondence between an input image and a reference face image is known, which can be determined by facial feature detection. The input image can be separated into texture and shape using a 2D face shape model [10]. The shape of a face is coded as the displacement field from the reference image, and the texture denotes an intensity map, which is produced by mapping the original image on to the reference image. All texture images have the same shape as the reference image. According to the authors, uneven illumination compensation is done on the texture image in order to avoid disturbing the shape information on the original image, i.e., in (1), all identities have the same surface normal direction distribution $\mathbf{n}$. After illumination compensation, the

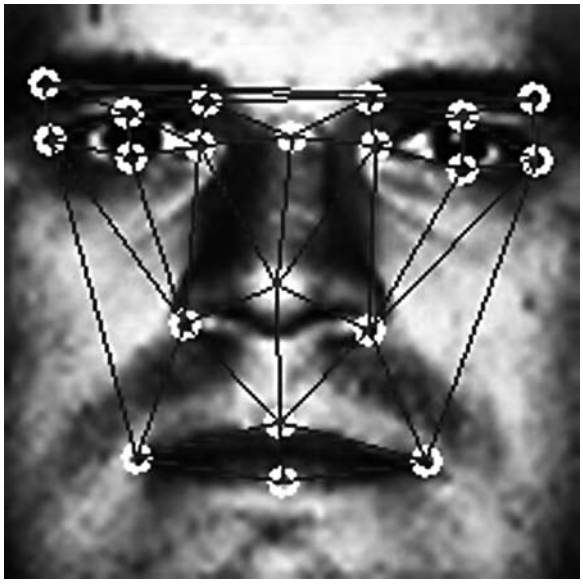
compensated texture and the original shape are combined to obtain the reconstructed image.

It is a challenge to find the pixel-wise correspondence between two pictures, especially when they are under uneven lighting conditions. Here, the position of some facial feature points, such as the eyebrows, eyes, nose, and mouth, are first determined manually (Fig. 2). The displacements of these key points between a facial image and the reference image are subsequently computed. The reference shape is obtained from the average 10 size-normalized and aligned images from the YaleB database. Using a triangle-based cubic interpolation method [11], the input image is mapped to the reference shape model. After processing the mapped texture, it can be mapped backwards from the reference shape to that of the original shape.

Lighting compensation is performed based on the mapped texture, which has a normal shape. From (4), the following is obtained

$$I(x,y) = \frac{I'(x,y) - B(x,y)}{A(x,y)}. \tag{5}$$

In order to avoid overflowing, all the intensity values of $I(x, y)$ are restricted to the range of [0, 255], so (5) can be rewritten as follows.



Illumination Compensation. **Figure 2** Facial feature points that are used to build a pixel-wise correspondence.

$$I(x,y) = \begin{cases} 0, & I(x,y) < 0, \\ 255, & I(x,y) > 255, \\ \dfrac{I'(x,y) - B(x,y)}{A(x,y)}, & \text{otherwise.} \end{cases} \tag{6}$$

Therefore, for an input image, if the illumination category is determined and the corresponding values of $A(x, y)$ and $B(x, y)$ are precomputed, (6) can be used to calculate the reconstructed image, which is under normal lighting conditions.

Besides the effect of illumination on appearance, face images of distinct subjects actually look quite different. This is because the appearance of a human face is also dependent on other factors, such as gender, race, and makeup. In order to accurately estimate the light source category, the distinctive elements of a person's appearance have to be eliminated to the extent possible while keeping the illumination information unchanged. Then, the illumination map is used to determine the illumination category. An image processed by the BHE [7] is considered as a reference image. The BHE processed image is then compared to the same image processed by the HE to obtain a pixel-wise difference between the two images. This difference image, which is called an illumination map, reflects the effect of the light source on different parts of the face image, and can be used to estimate the illumination category.

The determination of the illumination category is done by (LDA) ▶ *Linear Discriminant Analysis* [12]. The training images are divided into 65 different categories on the basis of their lighting conditions, and each category includes nine images that are under the same lighting condition and that belong to different people in the YaleB database.

Based on the training images in the YaleB database, the optimal values for $A_i(x, y)$ and $B_i(x, y)$ for each illumination category can be estimated by means of the least-squared method. For each illumination category i, suppose that the number of training samples equals m, then (4) can be rewritten as follows:

$$\begin{bmatrix} I'_1(x,y) \\ \vdots \\ I'_k(x,y) \\ \vdots \\ I'_m(x,y) \end{bmatrix} = \begin{bmatrix} I_1(x,y) & 1 \\ \vdots & \vdots \\ I_k(x,y) & 1 \\ \vdots & \vdots \\ I_m(x,y) & 1 \end{bmatrix} \begin{bmatrix} A_i(x,y) \\ B_i(x,y) \end{bmatrix}, \tag{7}$$

where $i = 1, \ldots, 65$ and $k = 1, \ldots, m$.

Let $F' = [I'_1(x,y) \cdots I'_k(x,y) \cdots I'_m(x,y)]^T$, where T represents the transpose, $I'_k(x,y)$ is the *kth* subject under the *ith* lighting category in the training set, and

$$F = \begin{bmatrix} I_1(x,y) & 1 \\ \vdots & \vdots \\ I_k(x,y) & 1 \\ \vdots & \vdots \\ I_m(x,y) & 1 \end{bmatrix},$$

where $I_k(x, y)$ represents the face of the *kth* subject in the training set under normal lighting conditions. Then (7) can be written as follows:

$$F' = F \begin{bmatrix} A_i(x, y) \\ B_i(x, y) \end{bmatrix}, \quad i = 1, \ldots, 65. \tag{8}$$

As the images in the different row of F, i.e., $I_k(x, y)$, are images of different people, they are independent of each other. The least-squared solution to (8) can be calculated as follows:

$$\begin{bmatrix} A_i(x, y) \\ B_i(x, y) \end{bmatrix} = \left(F^T F\right)^{-1} F^T F', \quad i = 1, \ldots, 65. \tag{9}$$

Using (9), the optimal value of $A_i(x, y)$ and $B_i(x, y)$ for the *ith* lighting category can be computed, and $A_i(x, y)$ and $B_i(x, y)$ are called A-map and B-map, respectively. (For more detailed descriptions, refer to Xie and Lam [7]).

Experimental Results

The training of the illumination compensation algorithm is based on the YaleB database. Therefore, the performance of the algorithm can be evaluated by utilizing other databases such as the Yale face database, [13] the YaleB face database, and the AR face database [14]. For each database, only images with an upright frontal view and a neutral expression are selected. The original images in the databases are shown in the first row of Figs. 3–5, images processed by the HE in the second row, those processed by the BHE in the third row, and images processed by the introduced algorithm in the fourth row.

After illumination compensation, the reconstructed images are used for face recognition. The PCA is used to measure the recognition rates after processing the images by different illumination compensation techniques. In each database, one image for each subject with normal illumination is selected as a

Illumination Compensation. Figure 3 Some experimental results based on the YaleB database. (**a**) original images, (**b**) images processed by the HE, (**c**) images processed by the BHE, (**d**) images processed by the introduced algorithm.

Illumination Compensation. **Figure 4** Some experimental results based on the yale database. **(a)** original images, **(b)** images processed by the HE, **(c)** images processed by the BHE, **(d)** images processed by the introduced algorithm.

Illumination Compensation. **Figure 5** Some experimental results based on the AR database. **(a)** original images, **(b)** images processed by the HE, **(c)** images processed by the BHE, **(d)** images processed by the introduced algorithm.

training sample, and others form the testing set. The respective recognition rates based on the different databases are presented in Table 1. From the experimental results, it can be seen that:

1. When the testing image set includes images under varying illumination, the HE can be utilized to improve the recognition performance as compared to that without using any preprocessing procedure.

Illumination Compensation. Table 1 Face recognition results using deferent preprocessing methods

Recognition rate (%)	None	HE	BHE	The introduced method
YaleB	43.4	61.4	77.5	99.5
Yale	36.7	36.7	80.0	90.0
AR	25.9	37.7	71.3	81.8
Combined	30.1	32.2	60.0	92.7

However, the improvement is very slight in some cases, e.g., for the Yale database.

2. The BHE or the introduced algorithm can improve the recognition rates significantly. The BHE leads to an improvement varying from 29.9 to 45.4%, and in the case of the introduced algorithm it varies from 53.3 to 62.6%. In other words, these two methods are both useful in eliminating the effect of uneven illumination on face recognition. In addition, the introduced algorithm can achieve the best performance level of all the methods used in the experiment.
3. The BHE method is very simple and does not require any prior knowledge. Compared to the traditional local contrast enhancement methods [5], its computational burden is for less. The main reason for this is that all the pixels within a block are equalized in the process, rather than just a single pixel, as in the adaptive block enhancement method. Nevertheless, as in the case of the traditional local contrast enhancement methods, noise is amplified after this process.
4. If the images processed by the HE are adopted to estimate the illumination category, the corresponding recognition rates using the different databases will be lowered. This is because variations between the images are affected not only by the illumination, but also by other factors, such as age, gender, race, and makeup. The illumination map can eliminate the distinctive personal information to the extent possible, while keeping the illumination information unchanged. Therefore, the illumination category can be estimated more accurately, and a more suitable illumination mode is selected.

The reconstructed facial images using the introduced algorithm appear to be very natural, and display great visual improvement and lighting smoothness. The effect of uneven lighting, including shadows, is almost eliminated. However, if there are glasses or a mustache, which are not Lambertian surface, in an image, some side effects may be seen under some special light source models. For instance, the glasses may disappear or the mustache may become faint.

Summary

This essay discussed a model-based method, which can be used for illumination compensation in face recognition. For a query image, the illumination category is first evaluated, followed by shape normalization, then the corresponding lighting model is used to compensate for uneven illumination. Next, the reconstructed texture is mapped backwards from the reference shape to that of the original shape in order to build an image under normal illumination. This lighting compensation approach is not only useful for face recognition when the faces are under varying illumination, but can also be used for face reconstruction. More importantly, the images of a query input are not required for training. In the introduced algorithm, 2D face shape model is adopted in order to address the effect of different geometries or shapes of human faces. Therefore, a more reliable and exact reconstruction of the human face is possible, and the reconstructed face is under normal illumination and appears more natural visually. Experimental results revealed that preprocessing the faces using the lighting compensation algorithm greatly improves the recognition rate.

Related Entries

► Face Recognition, Overview

References

1. Adini, Y., Moses, Y., Ullman, S.: Face recognition: the problem of compensating for changes in illumination direction. IEEE T. Pattern Anal. **19**(7), 721–732 (1997)
2. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cognitive Neurosci. **3**, 71–86 (1991)
3. Bartlett, M.S., Movellan, J.R., Sejnowski, T.J.: Face recognition by independent component analysis. IEEE T. Neural Networ. **13**(6), 1450–1464 (2002)
4. Yale University [Online]. Available at: http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html
5. Pizer, S.M., Amburn, E.P.: Adaptive histogram equalization and its variations. Comput. Vision Graph. **39**, 355–368 (1987)

6. Zhu, J., Liu, B., Schwartz, S.C.: General illumination correction and its application to face normalization. In proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, pp. 133–136. Hong Kong, China, (2003)
7. Xie, X., Lam, K.M.: Face recognition under varying illumination based on a 2D face shape model. Pattern Recogn. **38**(2), 221–230 (2005)
8. Xie, X., Lam, K.M.: An efficient illumination normalization method for face recognition. Pattern Recogn. Lett. **27**(6), 609–617 (2006)
9. Zhao, J., Su, Y., Wang, D., Luo, S.: Illumination ratio image: synthesizing and recognition with varying illuminations. Pattern Recogn. Lett. **24**(15), 2703–2710 (2003)
10. Liu, C., Wechsler, H.: A shape- and texture-based enhanced fisher classifier for face recognition. IEEE T. Image Process. **10**(4), 598–608 (2001)
11. Goshtasby, A.: Piecewise cubic mapping functions for image registration. Pattern Recogn. **20**(5), 525–533 (1987)
12. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE T. Pattern Anal. **19**(7), 711–720 (1997)
13. Yale University [Online]. Available at: http://cvc.yale.edu/projects/yalefaces/yalefaces.html
14. Martinez, A.M., Benavente, R.: The AR face database, CVC technical report #24 (1998)

Illumination Normalization

▶ Illumination Compensation

Image Acquisition

▶ Image Formation

Image Capture

▶ Biometric Sample Acquisition

Image Classification

▶ Image Pattern Recognition

Image Enhancement

The use of computer algorithms to improve the quality of an image by giving it higher contrast or making it less blurred or less noisy.

▶ Fingerprint Recognition, Overview
▶ Skull, Forensic Evidence of

Image Formation

Xiaoming Peng
College of Automation, University of Electronic Science and Technology of China, Chengdu, Sichuan, China

Synonym

Image acquisition

Definition

Image formation is the process in which three-dimensional (3D) scene points are projected into two-dimensional (2D) image plane locations, both geometrically and optically. It involves two parts. The first part is the geometry (derived from camera models assumed in the imaging process) that determines where in the image plane the projection of a scene point will be located. The other part of image formation, related with radiometry, measures the brightness of a point in the image plane as a function of illumination and surface properties.

Introduction

Common visual images result from light intensity variations across a two-dimensional plane. However, light is not the only source used in imaging. To understand how vision might be modeled computationally and replicated on a computer, it is important to understand the image acquisition process. Also, understanding image formation is a prerequisite before one could solve some complex computer vision tasks, such as

► camera, ► calibration and shape from shading. There are many types of imaging devices, ranging from biological vision systems (e.g., animal eyes) to video cameras and medical imaging machines (e.g., a Magnetic Resonance Imaging (MRI) scanner). Generally, the mechanisms of image creation vary across different imaging systems and as a result, it is not possible for this entry to cover all these mechanisms. In fact, the author confines the discussion of image formation process to a widely-used type of camera, the television camera in the visual spectrum. The reader, who is interested in image formation of other types of imaging devices might be referred to [1] for medical image formation, [2] for synthetic aperture radar (SAR) image formation, [3] for infrared image formation, and [4] for acoustic image formation. The image formation process of a television camera involves two parts. The first part is the camera geometry that determines where the projection of a scene point in the image plane will be located; the other part measures the amount of received light energy in individual pixels as the result of interaction among various scene surface material and light sources.

The Perspective Projection Camera Geometry

Some commonly-used geometric camera models include the perspective camera model, the weak-perspective camera model, the affine camera model, and the orthographic camera model. Among them, the perspective camera model is the most popular. In particular, an idealized model that defines perspective projection is the pinhole camera model, in which rays of light pass through a "pinhole" and form an inverted image of the object on the image plane. The geometry of the device is depicted in Fig. 1.

In Fig. 1, the optical center **C** coincides with the pinhole; the dotted vertical line perpendicular to the image plane is the optical axis; the intersection of the optical axis with the image plane is the principal

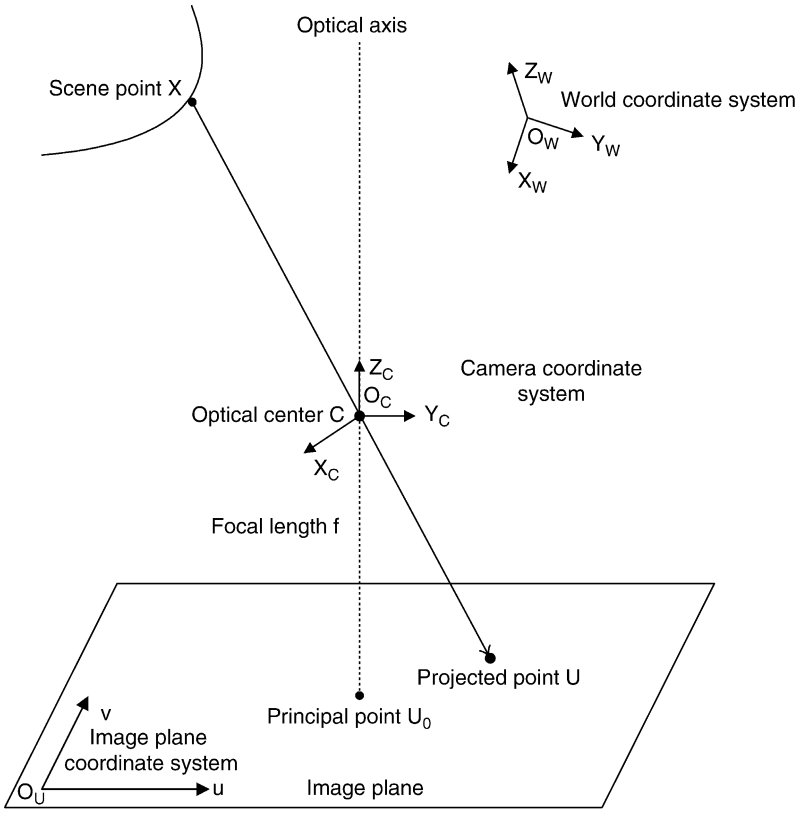


Image Formation. Figure 1 The perspective projection camera geometry.

point U_0; the focal length f is the distance between the optical center **C** and the principal point U_0. (It is assumed that the camera is focused at infinity.) The projection is formed by an optical ray, which is reflected from a 3D scene point $\mathbf{X} = (x, y, z)^T$ (where T denotes transpose) and passes through the optical center **C** and hits the 2D image plane at the point $\mathbf{U} = (u, v)^T$. Note that three coordinate systems are used in Fig. 1. The scene point **X** is expressed in the world coordinate system, which generally is not coincident with the camera coordinate system. The latter has the optical center **C** as its origin and its Z-axis is aligned with the optical axis. The alignment of these two coordinate systems can be performed by an Euclidean transformation consisting of a 3×3 rotation matrix **R** and a 3×1 translation vector, **t**. The projected point **U** on the image plane can be expressed in pixels in the image plane coordinate system, whose origin is assumed as the lower left corner of the image plane. Expressing $\mathbf{X} = (x, y, z)^T$ in the homogeneous form as $\tilde{\mathbf{X}} = (x, y, z, 1)^T$, the coordinates of **U** are given as [5]

$$\begin{cases} u = \frac{\mathbf{m}_1 \bullet \tilde{\mathbf{X}}}{\mathbf{m}_3 \bullet \tilde{\mathbf{X}}} \\ u = \frac{\mathbf{m}_2 \bullet \tilde{\mathbf{X}}}{\mathbf{m}_3 \bullet \tilde{\mathbf{X}}} \end{cases} \tag{1}$$

where $\mathbf{m}_1$, $\mathbf{m}_2$, and $\mathbf{m}_3$ are the three transposed rows of the 3×4 perspective projection matrix, and $\mathbf{M} = \mathbf{K}(\mathbf{R}/\mathbf{T})$ where "•" denotes the dot product operation.

The 3×3 matrix **K** in the perspective projection matrix **M** contains the intrinsic parameters of the camera and can be expressed as

$$\mathbf{K} = \begin{bmatrix} \alpha & -\alpha \cot \theta & u_0 \\ 0 & \frac{\beta}{\sin \theta} & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2}$$

where α and β are the scale factors of the image plane (in units of the focal length f), θ is the skew (it is the angle between the two image axis), and $(u_0, v_0)^T$ are the coordinates of the principal point. These five intrinsic parameters can be obtained through camera calibration.

It is worth pointing out that although the pinhole camera model is mathematically convenient, it is not the case with real cameras that are equipped with real lenses. The more realistic model of real lenses includes the radial distortion, a type of aberration that bends the light ray more or less than in the ideal case. As a consequence, the image formed is more or less distorted. Once the image is undistorted, the camera projection can be formulated as a projective projection.

The Radiometric Aspects that Determine the Brightness of an Image Point

A television camera measures the amount of received light energy in individual pixels as the result of interaction among various materials and various sources. The value measured is informally called brightness (or gray level). Radiometry is a branch of physics that deals with the measurement of the flow and transfer of radiant energy. It is an useful tool to establish the relationship between the brightness of a point in the image plane and the radiant energy the point received. Two definitions in radiometry, the radiance and irradiance, will be useful in the following explanations. Radiance is defined as the power of light that is emitted from an unit surface area into some spatial angle. The unit of radiance is W m^{-2} sr^{-1} (watts per square meter per steradian). Irradiance is defined as the amount of energy that an image-capturing device gets per unit of an efficient area of the camera. Gray-level of image pixels are quantized estimates of image irradiance. The unit of irradiance is W m^{-2} (watts per square meter).

Consider the relationship between the irradiance E measured in an infinitesimal image patch $\Delta\mathbf{I}$ and the radiance L produced by an infinitesimal scene patch $\Delta\mathbf{O}$ (Fig. 2). In Fig. 2, a lens with focal length f is placed at the coordinate origin and the infinitesimal scene patch $\Delta\mathbf{O}$ is at distance z from the lens. The off-axis angle spans between the optical axis and the line connecting $\Delta\mathbf{O}$ with $\Delta\mathbf{I}$. As given in [6],

$$E = L \frac{\pi}{4} \left(\frac{d}{f} \right)^2 \cos^4 \alpha, \tag{3}$$

where d is the diameter of the lens. Equation (3) shows that the image irradiance is proportional to the scene radiance.

As the image irradiance is a result of light reflection from scene objects, the ability of different materials to reflect light needs to be described. The most general model for this purpose is the bi-directional reflectance distribution function (BRDF), which describes the brightness of an elementary surface patch for a specific material, light source, and viewer directions. For most practically applicable surfaces, the BRDF remains

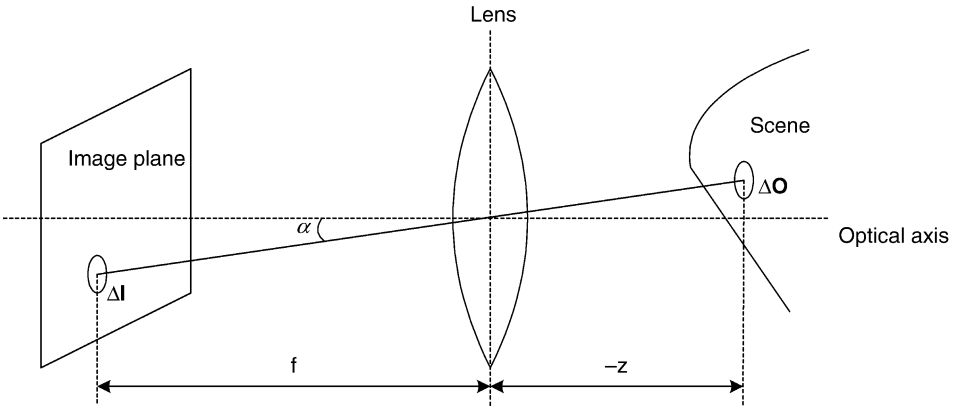


Image Formation. Figure 2 The relationship between image irradiance *E* and scene radiance *L*.

constant if the elementary surface patch rotates along its surface normal. For a Lambertian surface (also called ideal diffuse surface), which reflects light energy in all directions (and the radiance is constant in all directions) the BRDF is given as ρ/π, where ρ is the albedo of the surface.

Summary

Mechanisms of image formation vary across different types of imaging devices. The image formation process is explained both geometrically and optically of the television camera. The projected location of a 3D scene point onto the 2D image plane is dependent on the camera geometry, and the brightness of an image pixel is determined by the interaction among various material and various sources.

Related Entries

► Camera models
► Radiometric calibration

References

1. Webb, S.: The Physics of Medical Imaging. Taylor & Francis, London (1988)
2. Burns, B.L., Cordaro, J.T.: SAR image formation algorithm that compensates for the spatially variant effects of antenna motion. Proc. SPIE 2230, 14–24 (1994)
3. Foulkes, P.W.: Towards infrared image understanding. Ph.D. Thesis, Department of Engineering Science, Oxford University (1991)
4. Murino, V., Trucco, A.: A confidence-based approach to enhancing underwater acoustic image formation. IEEE Trans. Image Process. **8**(2), 270–285 (1999)
5. Forsyth, D.A., Ponce, J.: Computer Vision: A modern Approach. Prentice Hall, New Jersey (2003)
6. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision, 2nd edn. Brooks/Cole, California (1998)

Image Formation Process

► Face Sample Synthesis

Image Morphology

Morphological image processing is a theoretical model for digital image processing built upon lattice theory and topology. Morphological operators or filters rely only on the relative ordering of pixel values, not on their numerical values, and are suited especially to the processing of binary and grayscale images.

► Footwear Recognition

Image Pattern Classification

► Image Pattern Recognition

Image Pattern Recognition

Jian Yang, Jingyu Yang
School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, China

Synonyms

Image classification; Image pattern classification; Image recognition

Definition

Image pattern recognition is the problem of exploring how to recognize image patterns. An image pattern recognition system generally consists of four parts: a camera that acquires the image samples to be classified, an image preprocessor that improves the qualities of images, a feature extraction mechanism that gains discriminative features from images for recognition, and a classification scheme that classifies the image samples based on the extracted features.

Introduction

Image is the most important pattern perceived everyday. A lot of biometric patterns, such as faces, fingerprints, palmprints, hands, iris, ears, are all shown in images. Image pattern recognition, therefore, the fundamental problem in pattern recognition area, particularly in biometrics. The process of an image pattern recognition task generally includes four steps: image acquisition, image preprocessing, image feature extraction and classification, as shown in Fig. 1.

Image acquisition is the process of acquiring digital images from devices (e.g., digital cameras, scanners, or video frame grabbers) that are primarily used to capture still images. It is the elementary step of an image pattern recognition task. The quality of images acquired essentially affects the performance of the whole image pattern recognition system.

Image preprocessing is a common term for operations on images at the lowest level of abstraction. The objective is to improve the image data by removing the unwanted areas from the original image or enhancing some image features important for further processing. For biometric image recognition, image preprocessing generally involves two important aspects: image cropping and image enhancement.

Image feature extraction is the core problem of image pattern recognition, since finding most discriminative features of images is a key to solving pattern classification problems. Actually, our visual system has the special power of finding the most discriminative features. Given an image, its most important features can be obtained in a glance. Only based on these few features, the image can be immediately recognized next time. Note that for image pattern recognition, a direct way is to match two images based on pixel values. This method uses pixel values of an image as features, without a process of feature extraction. Generally speaking, this method has three weaknesses since image pattern is generally high-dimensional: first, it takes more time for classification; second, it increases the storage requirement of the recognition system; third, it may cause the so-called "curse of dimensionality" and achieve unsatisfactory recognition performance.

Classification is the task of classifying the image samples based on the extracted features and then providing the class label for images. A classifier needs to be designed or trained to do this task. There are a number of existing classifiers available for use. The choice of classifiers, however, is a difficult problem in practice. A more reliable way is to perform classifier combination, that is, to combine several classifiers to achieve a good, robust performance.

Image Preprocessing

In the context of biometrics, two image preprocessing steps deserve particular concern. One is *image cropping* and the other is *image enhancement*. Image cropping refers to the removal of the unwanted areas of an image to accentuate subject matter. For biometric images, the subject matter is the biometric object itself. For example, given a face image as shown in Fig. 2, the only concern is the face, since other parts, such as

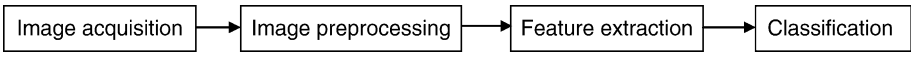


Image Pattern Recognition. **Figure 1** The general process of image pattern recognition.

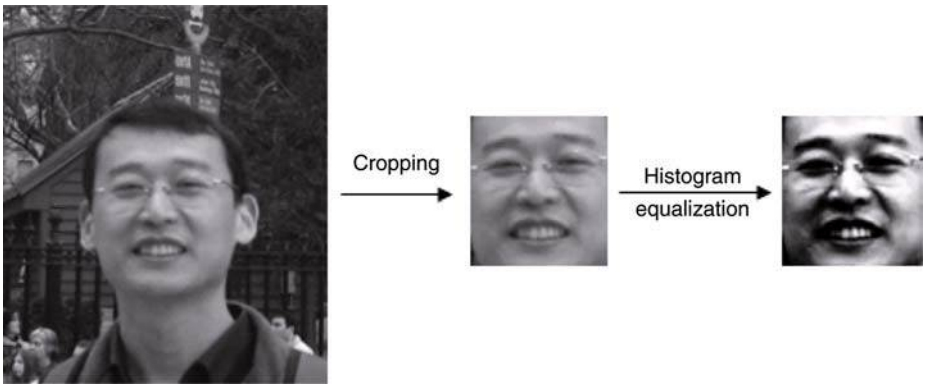


Image Pattern Recognition. Figure 2 Image cropping and histogram equalization.

clothes and background, are of little use face recognition. By proper cropping, the face is preserved for recognition. Figure 2 shows an example of image cropping. Image cropping can be done manually or automatically. Automatic image cropping is generally performed by virtue of an object detection algorithm, which is designed particularly for finding the object of interest.

Image enhancement is the improvement of digital image quality for visual inspection or for machine analysis, without knowing about the source of degradation. For biometric recognition, a primary role of image enhancement is to alleviate the effects of illumination. Image enhancement techniques can be divided into two broad categories: spatial domain methods which operate directly on pixels and frequency domain methods which operate on the ▶ Fourier transform of an image. A widely-used spatial domain method is ▶ histogram equalization, which aims to enhance the contrast of an image by using the normalized cumulative histogram as the grey scale mapping function [1]. Figure 2 shows a result of an image after histogram equalization. It should be noted that many image enhancement methods are problem-oriented: a method that works fine in one case may be completely inadequate for another case. For biometrics applications, an advantage for choosing image enhancement methods is that these methods can ultimately improve the recognition performance.

Image Feature Extraction

Broadly speaking, there are two categories of features to be extracted from an image: holistic features and local features. Image feature extraction generally involves a transformation from the input image space to the feature space. Holistic features are extracted through a holistic transformation, while the local features are derived via a local transformation.

Principal component analysis (PCA) and Fisher linear discriminant analysis (LDA) are two classical holistic transformation techniques. Both methods are widely applied to biometrics [2, 3]. However, PCA and LDA are essentially 1D vector pattern based techniques. Before applying PCA and LDA to 2D image patterns, the 2D image matrices must be mapped into 1D pattern vectors by concatenating their columns or rows. The pattern vectors generally lead to a high-dimensional space. For example, an image with a spatial resolution of 128 × 128 results in a 16,384-dimensional vector space. In such a high-dimensional vector space, computing the Eigenvectors of the covariance matrix is very time-consuming. Although the singular value decomposition (SVD) technique [2] is effective for reducing computations when the training sample size is much smaller than the dimensionality of the images, it does not help much when the training sample size becomes large.

Compared with PCA, the two-dimensional PCA method (2DPCA) [4] is a more efficient technique for dealing with 2D image pattern, as 2DPCA works on matrices rather than on vectors. Therefore, 2DPCA does not transform an image into a vector, but rather, it constructs an image covariance matrix directly from the original image matrices. In contrast to the covariance matrix of PCA, the size of the image covariance matrix of 2DPCA is much smaller. For example, if the image size is 128 × 128, the image covariance matrix of 2DPCA is still 128 × 128, regardless of the training sample size. As a result, 2DPCA has a remarkable computational advantage over PCA. In addition,

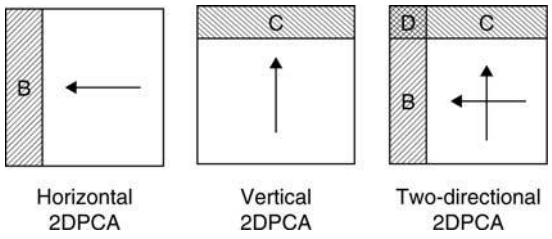


Image Pattern Recognition. **Figure 3** Illustration of the horizontal 2DPCA transform, the vertical 2DPCA transform, and the 2DPCA transform in both directions.

2DPCA can be performed on images horizontally and vertically. The horizontal 2DPCA is invariant to vertical image translations and vertical mirror imaging, and the vertical 2DPCA is invariant to horizontal image translations and horizontal mirror imaging. 2DPCA is therefore less sensitive to imprecise object detection and image cropping, and can improve the performance of PCA for image recognition [5]. Following the derivation of 2DPCA, some two-dimensional LDA (2DLDA) versions were developed and used for image feature extraction [6, 7] (Fig. 3).

Gabor wavelet transformation [8] and local binary patterns (LBP) [9] are two representative local transformation techniques. A Gabor wavelet (filter) is defined by a two-dimensional Gabor function, which was proposed by Daugman to model the spatial summation properties of the receptive fields of simple cells in the visual cortex [8]. A bank of Gabor filters with various scales and rotations are convolved with an image, resulting in a set of Gabor features of the image. Gabor features were demonstrated very effective for biometric image recognition due to their insensitivity to illumination variations. However, Gabor feature space is generally very high-dimensional, since multiple scales and rotations generate many filters. A feature extraction method, for example PCA or LDA, is usually followed to further reduce the dimension of the original Gabor feature space. In contrast to Gabor wavelet transformation, LBP is a relative new local descriptor method. This method commonly combines with a spatially enhanced histogram for image feature extraction [10].

Classification

After image feature extraction, one (or multiple) classifier for classification needs to be designed or chosen. The nearest neighbor (1-NN) classifier is one of the most widely-used classifiers due to its simplicity and effectiveness. Cover and Hart laid the theoretical foundation of 1-NN classifier and showed in 1967 [11] that when the training sample size approaches to infinity, the error rate of the NN classifier is bounded above by twice the Bayes error rate. As a generalization of 1-NN classifier, K-NN classifier was presented subsequently [12]. In practice, the performance of the NN-classifier depends on the representational capacity of prototypes as well as on how many prototypes are available. To enhance the representational capacity of the available limited prototypes, a nearest feature line classifier was proposed and applied to face biometrics [13].

Support vector machine (SVM) [14] is very popular as a classification method over the last decade. An SVM seeks to construct a separating hyperplane that maximizes the margin between the two classes of data sets. The margin is the perpendicular distance between the two parallel hyperplanes, each of which is determined by the class sample points closest to the separating hyperplane (these sample points are called support vectors). Intuitively, the larger the margin the better the ▶ generalization error of the classifier.

No classifier is perfect; each classifier has its advantages and weaknesses. To achieve a good, robust classification performance, multiple classifiers can be combined in practice. If the set of classifiers is determined, the remaining problem is to design or choose a proper combination scheme. A large number of classifier combination schemes (including voting, sum rule, bagging, boasting, etc.) have been proposed and summarized in the literature [15].

Summary

This entry provides a general view of image pattern recognition. The process of image pattern recognition includes four steps: image acquisition, image preprocessing, image feature extraction and classification. For image preprocessing, more attention is given to image cropping and image enhancement, which are two important steps in dealing with biometric images. For image feature extraction, holistic and local feature extraction methods are outlined, with an emphasis on the 2DPCA method which was designed to fit the two-dimensional image patterns. For classification,

the nearest neighbor classifier, support vector machine, and classifier combination methods have been introduced.

Related Entries

- ▶ Dimensionality Reduction
- ▶ Feature Extraction
- ▶ Image Formation
- ▶ Local Feature Filters

References

1. Russ, J.C.: The Image Processing Handbook (4th edn.). Baker & Taylor Press, Charlotte, NC (2002)
2. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
3. Belhumeur, P.N., Hespanha, J.P., Kriengman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell. **19**(7), 711–720 (1997)
4. Yang, J., Zhang, D., Frangi, A.F., Yang, J.-y.: Two-dimensional PCA: a new approach to face representation and recognition. IEEE Trans. Pattern Anal. Machine Intell. **26**(1), 131–137 (2004)
5. Yang, J., Liu, C.: Horizontal and vertical 2DPCA-based discriminant analysis for face verification on a large-scale database. IEEE Trans. Inf. Forensics Security. **2**(4), 781–792 (2007)
6. Yang, J., Yang, J.-y., Frangi, A.F., Zhang, D.: Uncorrelated projection discriminant analysis and its application to face image feature extraction. Intern. J. Pattern Recognit. Artif. Intell. **17**(8), 1325–1347 (2003)
7. Ye, J., Janardan, R., Li, Q.: Two-dimensional linear discriminant analysis. In: Proceedings of the Annual Conference on Advances in Neural Information Processing Systems (NIPS'04), Vancouver, BC, Canada (2004)
8. Daugman, J.G.: Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A. **2**, 1160–1169 (1985)
9. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: application to face recognition. IEEE Trans. Pattern Anal. Machine Intell. **28**(12), 2037–2041 (2006)
10. Liu, C., Wechsler, H.: Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Trans. Image Process. **11**(4), 467–476 (2002)
11. Cover, T.M., Hart, P.E.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theory. **13**(1), 21–27 (1967)
12. Fukunaga, K.: Introduction to Statistical Pattern Recognition (2nd edn.). Academic Press, New York (1990)
13. Li, S.Z., Lu, J.: Face recognition using the nearest feature line method. IEEE Trans. Neural Netw. **10**(2), 439–443 (1999)
14. Burge, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Min. Knowl. Disc. **2**, 121–167 (1998)
15. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: a review. IEEE Trans. Pattern Anal. Machine Intell. **22**(1), 4–37 (2000)
16. Generalization error. http://en.wikipedia.org/wiki/Generalization_error

Image Recognition

- ▶ Image Pattern Recognition

Image Regeneration from Templates

- ▶ Template Security

Image Resolution

Image resolution describes the detail an image holds. Higher resolution means more image detail. When applied to digital images, this term usually means *pixel resolution*, which is a pixel count in digital imaging.

- ▶ Skin Texture

Image Segmentation

Image segmentation is a technique that partitions a given input image into several different regions, and each region shares common features, like similar texture, similar colors or semantically reasonable object. The goal of image segmentation is to map every pixel of image to a group that is meaningful to human. Image segmentation is, most of the time, used to locate the object of interest, and find boundaries between objects.

The partitioning of pixels in an image into sets based on a common characteristics.

- ▶ Deformable Models
- ▶ Gait Recognition, Silhouette-Based
- ▶ Hand shape
- ▶ Iris Super-Resolution

Image Warping

The process of manipulating an image is such that the pixels of the original image are moved to new locations, without changing their color and intensity. In face processing, images are sometimes warped such that their shape is normalized, i.e., all warped images have the same width, height, eye location, etc.

- ▶ Face Alignment
- ▶ Face Device
- ▶ Face Tracking
- ▶ Face Recognition, Component-Based

Imaging Spectroscopy

- ▶ Multispectral and Hyperspectral Biometrics

Imaging Volume

The width of the imaging volume is defined by the field of view of the imaging system, and the depth of the imaging volume is defined by the depth of field of the imaging system.

- ▶ Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition

Implementation Under Test (IUT)

The implementation under test is that which implements the base standard(s) being tested. Depending on the conformance requirements of the base standard, this may simply be a set of biometric data interchange records (BDIRs) or it may be a computer algorithm or other product that creates the BDIRs and/or uses the data contained in the BDIRs.

- ▶ Conformance Testing for Biometric Data Interchange Formats, Standardization of

Impostor

A generic term for a person unknown by a biometric system who wishes to obtain the privileges of a client by claiming her/his identity.

- ▶ Liveness Detection: Iris
- ▶ Multiple Experts

Imposter Distribution

The probability distribution of the match score of a biometric for cases where two instances of biometric templates that are derived from different individuals are compared.

- ▶ Iris on the Move™

Impostor Match

Imposter match is the match between a pair of biometrics from two different persons.

- ▶ Fingerprint Matching, Automatic
- ▶ Individuality of Fingerprints

Imprecise Localization

The fiducial points on the face (such as eye corners, eye centers, eyebrows) cannot always be manually marked or automatically detected with pixel precision. This is generally due to the ambiguous nature of the gray or color patterns at these positions. Also, changes in illumination and pose, effect the perception of image rendering of such fiducials.

► Face Misalignment Problem
► Face Recognition, Component-Based

Incremental Learning

Xin Geng, Kate Smith-Miles
Deakin University, Melbourne, VIC, Australia

Synonyms

Adaptive learning; Online learning; Transfer learning

Definition

Incremental learning is a machine learning paradigm where the learning process takes place whenever new example(s) emerge and adjusts what has been learned according to the new example(s). The most prominent difference of incremental learning from traditional machine learning is that it does not assume the availability of a sufficient training set before the learning process, but the training examples appear over time.

Introduction

For a long time in the history of machine leaning, there has been an implicit assumption that a "good" training set in a domain is available a priori. The training set is so "good" that it contains all necessary knowledge that once learned, can be reliably applied to any new examples in the domain. Consequently, emphasis is put on learning as much as possible from a fixed training set. Unfortunately, many real-world applications cannot match this ideal case, such as in dynamic control systems, web mining, and time series analysis, where the training examples are often fed to the learning algorithms over time, i.e., the learning process is *incremental*. There are several reasons for the need for incremental learning:

1. It might be infeasible in time, storage, or other costs to obtain a sufficiently large number of representative examples before the learning process.
2. Even when ► training data, sufficiency can be obtained before learning, the learning algorithm might get computationally intractable if directly applied to all the available training data, or the whole training set cannot be loaded into the main memory.
3. When new examples become available, learning from scratch might waste time and computation resource. Instead, modifying learned knowledge according to the new examples might be a better choice, especially for those applications requiring real-time response.
4. If the example generation itself is time-dependent, e.g., ► time series data, then it inherently suits an incremental style of learning.
5. Nonstationary environments might change the target concept over time (► concept drift). In such case, the learner should be able to self-adapt to the changing environments.

In actual fact, incremental learning is quite common in reality. Some researchers even claim that incrementality is rather ubiquitous in learning [1], which can be evidenced by the way humans acquire knowledge over time. Although in some cases, such as theory refinement [2], all of the "teachable" knowledge may be available a priori, most learning tasks are inherently incremental. Interestingly, sometimes learning is only possible when data is presented incrementally. For example, Elman [3] ever gave an example of learning grammar with a recurrent network, where "the network fails to learn the task when the entire data set is presented all at once, but succeeds when the data are presented incrementally."

Considering the purpose of dealing with incrementality, several terms other than "incremental learning" has been used for the similar meanings. Some researchers [4] named the algorithms which can learn from increasing training examples as *online learning*

algorithms. The algorithms attempting to solve the concept drift problem are sometimes called *adaptive learning* algorithms [5], and some others are called *transfer learning* algorithms [6].

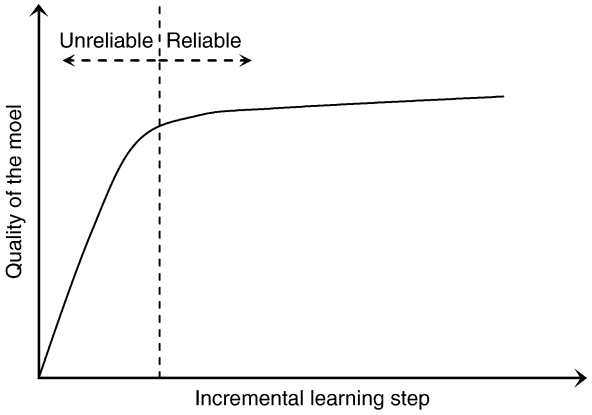
The possibly earliest incremental learning algorithm is the nearest neighbor classification method [7], although the term "incremental learning" was not explicitly proposed then. Most work on incremental learning starts from late 1980s. For example, Schlimmer and Fisher [8] proposed an algorithm named ID4 which incrementally generated a decision tree by updating the splits that were no longer the best given new examples. Aha et al. [9] proposed a framework called instance-based learning to solve incremental learning tasks using only specific instances. Syed et al. [10] found that the support vectors in SVM could form a succinct and sufficient training set for incremental learning. Ross et al. [11] proposed an incremental learning algorithm for visual tracking based on a low-dimensional subspace representation, which can well handle changes in the appearance of the target.

Evaluation Criteria

Same with traditional machine learning methods, many incremental learning algorithms were evaluated by the prediction accuracy on some benchmark data sets. Additional to this general criterion, researchers also proposed other more specific criteria according to the characteristics of incremental learning.

Schlimmer and Granger [12] proposed three criteria to measure the usefulness and effectiveness of an incremental learning method: (1) the number of observations (examples) needed to obtain a 'stable' concept description, (2) the cost of updating memory, and (3) the quality of learned concept descriptions. These measures are specifically designed for online algorithms trained on time-based examples.

More generally, Syed et al. [10] claimed that in order to measure an incremental learning algorithm, two new questions need to be answered: (1) How much better is a learned model at step $n + i$ than another model obtained after step n? (2) Can an incremental algorithm recover in the next incremental step(s), if it goes drastically off the "actual" concept at any stage? Consequently, they proposed three criteria for evaluation of the robustness of incremental learning algorithms: (1) *Stability* – the prediction accuracy on the test set should not vary wildly at every incremental learning step; (2) *Improvement* – there should be improvement in the prediction accuracy as the training progresses and the learning algorithm sees more training examples; and (3) *Recoverability* – the learning method should be able to recover from its errors, i.e., even if the performance drops at a certain learning step, the algorithm should be able to recover to the previous best performance.



Incremental Learning. Figure 1 A typical example of incremental learning curve.

Another often used criterion is the learning curve. An incremental algorithm may start learning from scratch and gradually obtain knowledge with an increasing amount of training examples. Consequently, the quality of the learned model (usually measured by prediction accuracy) displays a gradually improving curve over time, which is called a learning curve. A typical example of the learning curve of an incremental learning algorithm is shown in Fig. 1. Usually the learned model is not very reliable at the early stage of the curve. Decisions can be made according to the learning curve on how valuable the output of the incremental learner might be at a certain stage. However, in practice, it is often difficult to determine the point at which the model has learned "enough" to be reliable. Generally, a typical "good" learning curve should increase rapidly to a relatively steady high level.

Incremental Learning Tasks and Algorithms

The term incremental has been applied to both learning tasks and learning algorithms. Giraud–Carrier

[1] gave definition of incremental learning tasks and algorithms as follows:

Definition 1: A learning task is incremental if the training examples used to solve it become available over time, usually one at a time.

Definition 2: A learning algorithm is incremental if, for any given training sample $e_1,\ldots,e_n$, it produces a sequence of hypotheses $h_0,h_1,\ldots,h_n$, such that h_{i+1} depends only on h_i and the current example e_i.

This definition is based on the incrementality of training examples. Zhou and Chen [13] later extended this definition by distinguishing three different kinds of incremental tasks:

1. *Example-Incremental Learning Tasks (E-IL Tasks)*: New training examples are provided after a learning system is trained. For example, a face recognition system can gradually improve its accuracy by incorporating new face images of the registered users when they use it without reconfiguring and/or retraining the entire system. The description of the E-IL tasks is similar to Definition 1.
2. *Class-Incremental Learning Tasks (C-IL Tasks)*: New output classes are provided after a learning system is trained. For example, if a new user is added into the registered user group in the aforementioned face recognition system, the system should be able to recognize the new user without reconfiguring and/or retraining the entire system.
3. *Attribute-Incremental Learning Tasks (A-IL Tasks)*: New input attributes are provided after a learning system is trained. For example, if the camera used in the aforementioned face recognition system is changed from gray-scale camera to color camera, the system should be able to utilize the additional color features without reconfiguring and/or retraining the entire system.

While this taxonomy provides more insights into incremental learning, the definition based on incremental training examples has been widely accepted. In fact, new classes and new attributes are regarded as possible changes of the new examples, then C-IL and A-IL tasks can also be regarded as E-IL tasks. This is why we still use the incrementality of training examples to give definition of incremental learning at the beginning of this essay.

The relationship between incremental/nonincremental tasks and learners is described by two matrices in [1]: the application matrix and the utility matrix, as shown in Fig. 2. The application matrix indicates whether a particular learner can be applied to a particular task, while the utility matrix indicates whether a particular match of learner and task is of high utility. From the definitions of incremental learning task and algorithm (learner), it is natural to expect high utility of applying incremental learners to incremental tasks and nonincremental learners to nonincremental tasks. The off-diagonal match in the application matrix (incremental learner to nonincremental task, and nonincremental learner to incremental task) is also possible, but of low utility.

On the one hand, incremental learners can be applied to nonincremental tasks if taking the training examples one by one, although all of the training examples are available a priori. However, since incremental learners can only make use of current hypothesis and example, its "vision" is inherently local. In this case, the global information contained in the entire training set of the nonincremental task might be ignored by the incremental learner.

On the other hand, nonincremental learners can be applied to incremental tasks if retraining the system whenever a new example is available. This brute-force way is certainly inefficient or sometimes infeasible in terms of memory requirement (for the storage of all previous training samples) and computational resources. Moreover, in cases where concept drift occurs, the most recent examples provide most information for the new concept. Thus retraining on

a

		Learner	
		Incr.	Non-incr.
Task	Incr.	Yes	Yes
	Non-incr.	Yes	Yes

b

		Learner	
		Incr.	Non-incr.
Task	Incr.	High	Low
	Non-incr.	Low	High

Incremental Learning. **Figure 2** Relationship between incremental/nonincremental tasks and learners [1]: (**a**) application matrix and (**b**) utility matrix.

the entire training set might not be able to perceive the changes of the target concept.

Even when applied to incremental tasks, the suitable incremental learning algorithm must be carefully designed. In cases where the target concept is relatively stable or gradually changes, the new examples help to *refine* the existing learned concept model. In cases where the concept changes are substantial, the system might need to discount or even "forget" the old examples, and *adjust* what has been induced from them. In some applications, the incremental learning algorithms even have to automatically distinguish the above two cases and take actions to either improve the current concept representation as more supportive examples become available, or respond to suspected changes in the definition of the target concept when the incoming examples become inconsistent with the learned concept.

Applications

With increasing demands from real applications for machine learning algorithms to be more adaptive, scalable, robust, and responsive, incremental learning has been successfully applied to solve a wide range of real-world problems. Generally speaking, incremental learning is most suitable for the following three kinds of applications:

1. Applications where the target concepts change over time.
   Examples:
   (a) *Robotics*. The environment around a robot is often changing and unpredictable. In order to accomplish the assigned missions, a robot must be able to adapt to the new environment and react properly, which is naturally an incremental learning process.
   (b) *Intelligent agent*. An intelligent agent is an entity which can observe and act upon an environment and direct its activity towards achieving goals. Incremental learning can help an intelligent agent to perceive changes of the environment and accordingly adjust its strategy to achieve goals.
2. Applications where the "sufficient training sets" are too big.
   Examples:
   (a) *Content based image retrieval (CBIR)*. CBIR is the problem of searching for digital images in large databases or from the internet based on analysis of the actual contents of the images. In order to learn sufficient concepts for retrieval purpose, usually a large amount of images should be included in the training set. The size of the training set could even be infinite in the case of online image retrieval. Incremental learning can be used to solve the problem of shortage in computation and storage resources. Also it can help to implement an "improve while using" system by gradually improving accuracy whenever new examples emerge during the use of the system.
   (b) *Face recognition*. The main challenge of automatic face recognition is that face images could present changes in various aspects, such as pose, illumination, expression, and occlusion. Taking all of these affective factors into consideration consequently requires a huge training set. Similarly, incremental learning techniques can help to realize a face recognition system which learns while in use, i.e., every time a user uses the system, a new face image is provided to the system for incremental learning.
3. Applications where the training examples are obtained over time (time series data).
   Examples:
   (a) *Visual object tracking*. During visual tracking, the appearance of the target object is usually highly variable (e.g., pose variation, shape deformation, illumination changes, etc.) in different video frames over time. Incremental learning techniques can be adopted to update the internal representation of the target object in realtime for constantly and efficiently tracking the object.
   (b) *Software project estimation*. Estimating the cost and duration of a software project is very important for project management. However, useful information about a project becomes available over time while the project progresses. Thus software project estimation is inherently an incremental learning task.

Summary

Incrementality is part of the nature of learning. Compared with traditional machine learning which requires

a training set beforehand, incremental learning shows several advantages: (1) It does not require a sufficient training set before learning; (2) It can continuously learn to improve when the system is running; (3) It can adapt to changes of the target concept; (4) It requires less computation and storage resources than learning from scratch; (5) It naturally matches the applications depending on time series. Nevertheless, incremental learning is not suitable for many non-incremental learning tasks due to the fact that it is inherently "myopic" and tends to ignore the global information in the entire training set.

Related Entries

► Machine-Learning

References

1. Giraud-Carrier, C.G.: A note on the utility of incremental learning. AI Commun. **13**(4), 215–224 (2000)
2. Ourston, D., Mooney, R.J.: Theory refinement combining analytical and empirical methods. Artif. Intell. **66**(2), 273–309 (1994)
3. Elman, J.L.: Learning and development in neural networks: The importance of starting small. Cognition **46**(1), 71–99 (1993)
4. Cheng, L., Vishwanathan, S.V.N., Schuurmans, D., Wang, S., Caelli, T.: Implicit online learning with kernels. In: Advances in Neural Information Processing Systems 19, pp. 249–256. Vancouver, Canada (2006)
5. Huo, Q., Lee, C.H.: On-line adaptive learning of the continuous density hidden markov model based on approximate recursive bayes estimate. IEEE Trans. Speech Audio Process. **5**(2), 161–172 (1997)
6. Pan, S.J., Kwok, J.T., Yang, Q.: Transfer learning via dimensionality reduction. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 677–682. Chicago, IL (2008)
7. Cover, T.M., Hart, P.E.: Nearest neighbour pattern classification. Trans. Inf. Theory **13**, 21–27 (1967)
8. Schlimmer, J.C., Fisher, D.H.: A case study of incremental concept induction. In: Proceedings of the National Conference on Artifical Intelligence, pp. 496–501. San Mateo, CA (1986)
9. Aha, D.W., Kibler, D.F., Albert, M.K.: Instance-based learning algorithms. Mach. Learn. **6**, 37–66 (1991)
10. Syed, N.A., Liu, H., Sung, K.K.: Handling concept drifts in incremental learning with support vector machines. In: Proceedings of ACM International Conference on Knowledge Discovery and Data Mining, pp. 317–321. San Diego, CA (1999)
11. Ross, D.A., Lim, J., Lin, R.S., Yang, M.H.: Incremental learning for robust visual tracking. Int. J. Comput. Vis. **77**(1-3), 125–141 (2008)
12. Schlimmer, J.C., Granger, R.H.: Incremental learning from noisy data. Mach. Learn. **1**(3), 317–354 (1986)
13. Zhou, Z.H., Chen, Z.: Hybrid decision tree. Knowl. Based Syst. **15**(8), 515–528 (2002)

Independent Component Analysis

Seungjin Choi
Department of Computer Science, Pohang University of Science and Technology, Korea

Synonyms

Blind source separation; Independent factor analysis

Definition

Independent component analysis (ICA) is a statistical method, the goal of which is to decompose multivariate data into a linear sum of non-orthogonal basis vectors with coefficients (encoding variables, latent variables, hidden variables) being statistically independent. ICA generalizes a widely-used subspace analysis method such as principal component analysis (PCA) and factor analysis, allowing latent variables to be non-Gaussian and basis vectors to be non-orthogonal in general. Thus, ICA is a density estimation method where a linear model is learnt such that the probability distribution of the observed data is best captured, while factor analysis aims at best modeling the covariance structure of the observed data.

Introduction

Linear latent variable model assumes that m-dimensional observed data $\boldsymbol{x}_t \in \mathbb{R}^m$ is generated by

$$\boldsymbol{x}_t = \boldsymbol{a}_1 s_{1,t} + \boldsymbol{a}_2 s_{2,t} + \cdots \boldsymbol{a}_n s_{n,t} + \epsilon_t, \qquad (1)$$

where $\boldsymbol{a}_i \in \mathbb{R}^m$ are *basis vectors* and $s_{i,t}$ are *latent variables* (hidden variables, coefficients, encoding variables) which are introduced for parsimonious representation ($n \leq m$). Modeling uncertainty or noise is absorbed in $\epsilon_t \in \mathbb{R}^m$. Neglecting the uncertainty ϵ_t in (1), the linear latent variable model is nothing but

linear transformation. For example, if $\boldsymbol{a}_i$ are chosen as Fourier or Wavelet basis vectors, then $s_{i,t}$ are associated Fourier or Wavelet coefficients that are served as *features* in pattern recognition. In the case where $\boldsymbol{a}_i$ are orthonormal, it is referred to as *orthogonal transformation*. Subspace analysis methods that are popular in pattern recognition also considers the linear model (1), assuming Gaussian factors $s_{i,t}$ and (isotropic) independent Gaussian noise ϵ_t. In such a case, model parameters such as $\boldsymbol{a}_i$ and diagonal covariance matrix of ϵ_t are estimated by expectation maximization algorithms [1].

The simplest form of independent component analysis (ICA) considers a noise-free linear latent variable model with assuming $m = n$, where observed variables $\boldsymbol{x}_t \in \mathbb{R}^n$ is assumed to be generated by

$$\boldsymbol{x}_t = \sum_{i=1}^{n} \boldsymbol{a}_i s_{i,t} = \boldsymbol{A}\boldsymbol{s}_t, \tag{2}$$

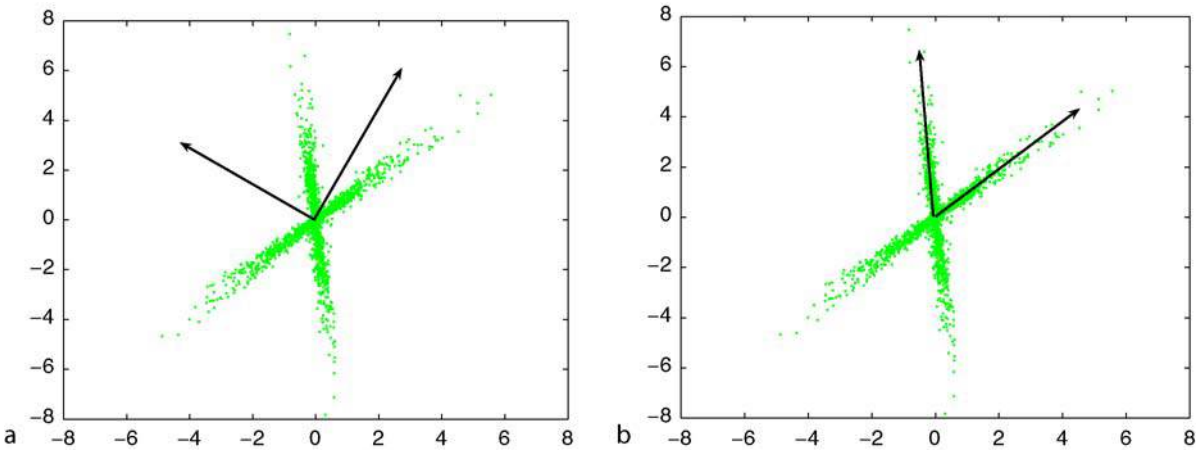
where $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n] \in \mathbb{R}^{m \times n}$ is referred to as *mixing matrix* and $\boldsymbol{s}_t = [s_{1,t}, \ldots, s_{n,t}]^\top \in \mathbb{R}^n$ are constrained to have *independent components*. ICA generalizes PCA in the sense that latent variables (components) are non-Gaussian and $\boldsymbol{A}$ is allowed to be non-orthogonal transformation, whereas PCA considers only orthogonal transformation and implicitly assumes Gaussian components. Fig. 1 shows a simple example, emphasizing the main difference between PCA and ICA. Exemplary basis face images learned by PCA and ICA are shown in Fig. 2.

Methods

The task of ICA is to estimate the mixing matrix $\boldsymbol{A}$ or its inverse $\boldsymbol{W} = \boldsymbol{A}^{-1}$ (referred to as *dexming matrix*) such that elements of the estimate $\boldsymbol{y}_t = \boldsymbol{A}^{-1}\boldsymbol{x}_t = \boldsymbol{W}\boldsymbol{x}_t$ are as independent as possible. A variety of methods for ICA have been developed so far. The following books are good resources for comprehensive understanding on ICA: Lee [2] where a unified view of several different principles, including ▶ mutual information minimization, information maximization, maximum likelihood estimation, and negentropy maximization are found; Hyvärinen et al. [3] where many useful fundamental background on ICA and FastICA algorithms are found; Cichocki and Amari [4] where various methods of source separation in the perspective of signal processing can be found. In addition to these books, several tutorial or review papers are also available [5, 6].

Methods for ICA can be categorized into two groups:

- *Unsupervised learning methods*: Factorial coding is a primary principle for efficient information representation and is closely related to



Independent Component Analysis. Figure 1 Two-dimensional data with two main arms are fitted by two different basis vectors: (**a**) PCA makes the implicit assumption that the data have a Gaussian distribution and determines the optimal basis vectors that are orthogonal, which are not efficient at representing non-orthogonal distributions; (**b**) ICA does not require that the basis vectors be orthogonal and considers non-Gaussian distributions, which is more suitable in fitting more general types of distributions.

Independent Component Analysis. **Figure 2** Sample face images from Yale DB are shown in (**a**). First 20 basis images determined by: (**b**) PCA; (**c**) ICA. A first application of ICA to face recognition is found in [25].

redundancy reduction that provides a principled method for unsupervised learning [7]. It is also related to ICA, aiming at a linear data representation that best model the probability distribution of the data, so higher-order statistical structure is incorporated.

– Maximum likelihood estimation (Kullback matching): It is well known that maximum likelihood estimation is equivalent to Kullback matching where the optimal model is estimated by minimizing Kullback-Leibler (KL) divergence between empirical distribution and model distribution. We consider KL divergence from the empirical distribution $\tilde{p}(\boldsymbol{x})$ to the model distribution $p_\theta(\boldsymbol{x})$

$$\begin{aligned} KL[\tilde{p}(\boldsymbol{x})||p_\theta(\boldsymbol{x})] &= \int \tilde{p}(\boldsymbol{x}) \log \frac{\tilde{p}(\boldsymbol{x})}{p_\theta(\boldsymbol{x})} d\boldsymbol{x} \\ &= -H(\tilde{p}) - \int \tilde{p}(\boldsymbol{x}) \log p_\theta(\boldsymbol{x}) d\boldsymbol{x}, \end{aligned} \tag{3}$$

where $H(\tilde{p}) = -\int \tilde{p}(\boldsymbol{x}) \log \tilde{p}(\boldsymbol{x}) d\boldsymbol{x}$ is the entropy of $\tilde{p}$. Given a set of data points, $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ drawn from the underlying distribution $p(\boldsymbol{x})$, the empirical distribution $\tilde{p}(\boldsymbol{x})$ puts probability $\frac{1}{N}$ on each data point, leading to

$$\tilde{p}(\boldsymbol{x}) = \frac{1}{N} \sum_{t=1}^{N} \delta(\boldsymbol{x} - \boldsymbol{x}_t). \tag{4}$$

It follows from (3) that

$$\arg_{\boldsymbol{\theta}} \min KL[\tilde{p}(\boldsymbol{x})||p_\theta(\boldsymbol{x})] \equiv \arg_{\boldsymbol{\theta}} \max \langle \log p_\theta(\boldsymbol{x}) \rangle_{\tilde{p}}, \tag{5}$$

where $\langle \cdot \rangle_{\tilde{p}}$ represents the expectation with respect to the distribution $\tilde{p}$. Plugging (4) into the right-hand side of (3), leads to

$$\begin{aligned} \langle \log p_\theta(\boldsymbol{x}) \rangle_{\tilde{p}} &= \frac{1}{N} \int \sum_{t=1} N \delta(\boldsymbol{x} - \boldsymbol{x}_t) \log p_\theta(\boldsymbol{x}) d\boldsymbol{x} \\ &= \frac{1}{N} \sum_{t=1}^{N} \log p_\theta(\boldsymbol{x}_t). \end{aligned} \tag{6}$$

Apart from the scaling factor $\frac{1}{N}$, this is just the log-likelihood function. In other words, maximum likelihood estimation is obtained from the minimization of (3).

– Mutual information minimization: Mutual information is a measure for statistical independence. Demixing matrix $\boldsymbol{W}$ is learned such that the mutual information of $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$ is minimized, leading to the following objective function:

$$\begin{aligned} \mathcal{J}_{mi} &= \int p(\boldsymbol{y}) \log \left[\frac{p(\boldsymbol{y})}{\prod_{i=1}^{n} p_i(y_i)} \right] d\boldsymbol{y} \\ &= -H(\boldsymbol{y}) - \left\langle \sum_{i=1}^{n} \log p_i(y_i) \right\rangle. \end{aligned} \tag{7}$$

Note that $p(\boldsymbol{y}) = \frac{p(\boldsymbol{x})}{|\det \boldsymbol{W}|}$. Thus, the objective function (7) is given by

$$\mathcal{J}_{mi} = -\log|\det \boldsymbol{W}| - \sum_{i=1}^{n} \langle \log p_i(y_i) \rangle, \tag{8}$$

where $\langle \log p(\boldsymbol{x}) \rangle$ is left out since it does not depend on parameters $\boldsymbol{W}$.

– Information maximization: Infomax [8] involves the maximization of the output entropy $\boldsymbol{z} = g(\boldsymbol{y})$ where $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$ and $g(\cdot)$ is a squashing function (e.g., $g_i(y_i) = \frac{1}{1+e^{-y_i}}$). It was shown that infomax contrast maximization is equivalent to the minimization of KL divergence between the distribution of $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$ and the distribution $p(\boldsymbol{s}) = \prod_{i=1}^{n} p_i(s_i)$. In fact, infomax is nothing but mutual information minimization in ICA framework.

– Negentropy maximization: Negative entropy or negentropy is a measure of distance to Gaussianity, yielding the larger value for random variable whose distribution is far from Gaussian. Negentropy is always nonnegative and vanishes if and only if the random variable is Gaussian. Negnetropy is defined as

$$J(\boldsymbol{y}) = H(\boldsymbol{y}^G) - H(\boldsymbol{y}), \tag{9}$$

where $\boldsymbol{y}^G$ is a Gaussian random vector whose mean vector and covariance matrix are the same as $\boldsymbol{y}$. It is shown that the negentropy maximization is equivalent to the mutual information minimization [2].

- *Algebraic methods*: Algebraic methods have been developed mainly for blind source separation (BSS), the task of which is to restore unknown sources $\boldsymbol{s}$ without the knowledge of $\boldsymbol{A}$, given the observed data $\boldsymbol{x}$. They, in general, are based on eigen-decomposition of certain statistical information matrix such as covariance matrix (or correlation matrix), higher-order moment matrix, or cumulant matrix.

– Generalized eigenvalue decomposition: Simultaneous diagonalization of two covariance matrices

with distinct eigenvalues achieves BSS. Earlier work includes FOBI [9] and AMUSE [10].

– Joint approximate diagonalization: Statistical efficiency increases when several covariance matrices or cumulant matrices are considered for joint approximate diagonalization. JADE [11] considers 4th-order cumulant matrices, SOBI [12] uses time-delayed correlation matrices, and SEONS [13] incorporates time-varying correlation matrices. All these methods use Jacobi rotation to jointly diagonalize the matrices considered. On the other hand, correlation matching is an alternative method, which is solved by least squares technique [14, 15].

Algorithms

Latent variables s_i or their estimates y_i are assumed to be statistically independent, i.e., the joint distribution is factored into the product of marginal distributions

$$p(\boldsymbol{s}) = \prod_{i=1}^{n} p_i(s_i), \quad or \quad p(\boldsymbol{y}) = \prod_{i=1}^{n} p_i(y_i). \tag{10}$$

It follows from the relation $p(\boldsymbol{x}) = p(\boldsymbol{s})/|\det \boldsymbol{A}|$ that the single factor of log-likelihood is given by

$$\log p_\theta(\boldsymbol{x}) = -\log|\det \boldsymbol{A}| + \sum_{i=1}^{n} \log p_i(s_i). \tag{11}$$

Then the objective function for on-line learning is given by

$$\mathcal{J} = -\log p_\theta(\boldsymbol{x}) = \log|\det \boldsymbol{A}| - \sum_{i=1}^{n} \log p_i(s_i), \tag{12}$$

which is equivalent to (8) that is used for mutual information minimization.

The gradient descent method gives a learning algorithm for $\boldsymbol{A}$ that has the form

$$\Delta \boldsymbol{A} = -\eta \frac{\partial \mathcal{J}}{\partial \boldsymbol{A}} = -\eta \boldsymbol{A}^{-\top}\{\boldsymbol{I} - \varphi(\boldsymbol{s})\boldsymbol{s}^\top\}, \tag{13}$$

where $\eta > 0$ is the learning rate and $\varphi(\boldsymbol{s})$ is the negative *score function* whose ith element $\varphi_i(s_i)$ is given by

$$\varphi_i(s_i) = -\frac{d \log p_i(s_i)}{ds_i}. \tag{14}$$

Employing the natural gradient [16], we have

$$\Delta \boldsymbol{A} = -\eta \boldsymbol{A}\boldsymbol{A}^\top \frac{\partial L}{\partial \boldsymbol{A}} = -\eta \boldsymbol{A}\{\boldsymbol{I} - \varphi(\boldsymbol{s})\boldsymbol{s}^\top\}. \tag{15}$$

At each iteration, latent variables $\boldsymbol{s}$ are computed by $\boldsymbol{s} = \boldsymbol{A}^{-1}\boldsymbol{x}$ using the current estimate of $\boldsymbol{A}$. Then the value of $\boldsymbol{A}$ is updated by (15). This procedure is repeated until $\boldsymbol{A}$ converges.

The function $\varphi_i(s_i)$ depends on the prior $p_i(s_i)$ that has to be specified in advance. Depending to the choice of prior, we have different data representation. In the case of Laplacian prior, the function $\varphi_i(s_i)$ has the form

$$\varphi_i(s_i) = sgn(s_i), \tag{16}$$

where sgn$(\cdot)$ is the signum function. Sparseness constraint was shown to be useful to describe the receptive field characteristics of simple cells in primary visual cortex [17]. Generalized Gaussian prior for s_i is useful in approximating most of uni-modal distributions [18].

Alternatively, it is possible to learn $\boldsymbol{A}^{-1}$ instead of $\boldsymbol{A}$. The $\boldsymbol{A}^{-1}$ coincides with the ICA filter [8]. If we define $\boldsymbol{W} = \boldsymbol{A}^{-1}$, then the natural gradient learning algorithm for $\boldsymbol{W}$ is given by

$$\Delta \boldsymbol{W} = \eta\{\boldsymbol{I} - \varphi(\boldsymbol{y})\boldsymbol{y}^\top\}\boldsymbol{W}. \tag{17}$$

This is a well-known ICA algorithm [19].

Softwares

We briefly introduce several ICA softwares so that one can immediately play with these MATALB codes or toolboxes to see how they are working on data sets. ICA Central (URL: http://www.tsi.enst.fr/icacentral/) was created in 1999 to promote research on ICA and blind source separation by means of public mailing lists, a repository of data sets, a repository of ICA/BSS algorithms, and so on. ICA Central might be the first place where you can find data sets and ICA algorithms. In addition, several widely-used softwares include

1. *ICALAB Toolboxes* (http://www.bsp.brain.riken.go.jp/ICALAB/): ICALAB is an ICA Matlab software toolbox developed in laboratory for Advanced Brain Signal Processing in RIKEN Brain Science Institute, Japan. It consists of two independent packages: ICALAB for signal processing and

ICALAB for image processing and each package contains a variety of algorithms.

2. *FastICA* (http://www.cis.hut.fi/projects/ica/fastica/): It is the FastICA Matlab package that implements fast fixed-point algorithms for non-Gaussianity maximization [3]. It was developed in Helsinki University of Technology, Finland and other environments (R, C++, Physon) are also available.
3. *Infomax ICA* (http://www.cnl.salk.edu/~tewon/ica_cnl.html): Matlab and C codes for Bell and Sejnowski's Infomax algorithm [8] and extended infomax [2] where, a parametric density model is incorporated into Infomax to handle both super-Gaussian and sub-Gaussian sources.
4. *EEGLAB* (http://sccn.ucsd.edu/eeglab/): EEGLAB is an interactive Matlab toolbox for processing continuous and event-related EEG, MEG and other electrophysiological data using ICA, time/frequency analysis, artifact rejection, and several modes of data visualization.
5. *ICA: DTU Toolbox* (http://isp.imm.dtu.dk/toolbox/ica/): 'ICA: DTU Toolbox' is a collection of ICA algorithms that includes: (1) 'icaML' which is an efficient implementation of Infomax; (2) 'icaMF' which is an iterative algorithm that offers a variety of possible source priors and mixing matrix constraints (e.g., positivity) and can also handle over and under-complete mixing; (3) 'icaMS' which is an 'one shot' fast algorithm that requires time correlation between samples.

Further Issues

1. *Overcomplete representation*: Overcomplete representation enforces the latent space dimension n to be greater than the data dimension m in the linear model (1). Sparseness constraints on latent variables are necessary to learn fruitful representation [20].
2. *Bayesian ICA*: In contrast to the simple ICA model (2), Bayesian ICA incorporates uncertainty and prior distributions of latent variables into the model (1). Independent factor analysis [21] is a pioneering work along this direction. EM algorithm for ICA was developed in [22] and full Bayesian learning was adopted [23].
3. *Kernel ICA*: Kernel methods were introduced to consider statistical independence in reproducing kernel Hilbert space [24], developing kernel ICA.
4. *Nonnegative ICA*: Nonnegativity constraints were imposed on latent variables, yielding nonnegative ICA [25]. Rectified Gaussian prior can also be used in Bayesian ICA to handle nonnegative latent variables. It was successfully applied in medical imaging [26].
5. *Nonstationarity*: Nonstationary characteristics was used in developing second-order source separation methods [27, 28].
6. *Sparseness*: Sparse component analysis is one of hot issues [29].
7. *Beyond ICA*: Independent subspace analysis [30] and tree-dependent component analysis [31] generalizes ICA, allowing intra-dependence structure in feature subspaces or clusters.

Summary

ICA has been successfully applied to various applications of pattern recognition. We have presented a brief overview of ICA, starting from the fundamental idea on learning a linear latent variable model for parsimonious representation. Natural gradient ICA algorithms were derived in the framework of maximum likelihood estimation or Kullback matching. Various softwares for ICA were introduced, so that one could easily apply ICA to his or her own applications. Further issues were also briefly mentioned so that readers could follow the status ICA.

References

1. Tipping, M.E., Bishop, C.M.: Mixtures of probabilistic principal component analyzers. Neural Comput. **11**(2), 443–482 (1999)
2. Lee, T.W.: Independent Component Analysis: Theory and Applications. Kluwer Academic: Boston (1998)
3. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. Wiley: New York (2001)
4. Cichocki, A., Amari, S.: Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. Wiley: West Sussex, England (2002)
5. Hyvärinen, A.: Survey on independent component analysis. Neural Computing Surveys **2**, 94–128 (1999)
6. Choi, S., Cichocki, A., Park, H.M., Lee, S.Y.: Blind source separation and independent component analysis. A review. Neural Information Processing - Letters and Review **6**(1), 1–57 (2005)
7. Barlow, H.B.: Unsupervised learning. Neural Computation **1**, 295–311 (1989)
8. Bell, A., Sejnowski, T.: An information maximisation approach to blind separation and blind deconvolution. Neural Comput. **7**, 1129–1159 (1995)

9. Cardoso, J.F.: Source separation using higher-order moments. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (1989)
10. Tong, L., Soon, V.C., Huang, Y.F., Liu, R.: AMUSE: a new blind identification alogrithm. In: Proceedings of the IEEE International Symposium on Circuits and Systems, pp. 1784–1787 (1990)
11. Cardoso, J.F., Souloumiac, A.: Blind beamforming for non Gaussian signals. IEE Proceedings-F **140**(6), 362–370 (1993)
12. Belouchrani, A., Abed-Merain, K., Cardoso, J.F., Moulines, E.: A blind source separation technique using second order statistics. IEEE Trans. Signal Processing **45**, 434–444 (1997)
13. Choi, S., Cichocki, A., Belouchrani, A.: Second order nonstationary source separation. J. VLST Signal Process Syst. Signal Image Video Technol. **32**, 93–104 (2002)
14. Choi, S., Cichocki, A., Belouchrani, A.: Blind separation of second-order nonstationary and temporally colored sources. In: Proceedings of IEEE Workshop on Statistical Signal Processing, pp. 444–447. Singapore (2001)
15. Choi, S., Cichocki, A.: Correlation matching approach to source sepration in the presence of spatially correlated noise. In: Proceedings of the IEEE International Symposium on Signal Processing and Applications. Kuala-Lumpur, Malaysia (2001)
16. Amari, S.: Natural gradient works efficiently in learning. Neural Comput. **10**(2), 251–276 (1998)
17. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature **381**, 607–609 (1996)
18. Choi, S., Cichocki, A., Amari, S.: Flexible independent component analysis. J. VLST Signal Process Syst. Signal Image Video Technol. **26**(1/2), 25–38 (2000)
19. Amari, S., Chen, T.P., Cichocki, A.: Stability analysis of learning algorithms for blind source separation. Neural Netw. **10**(8), 1345–1351 (1997)
20. Lewicki, M.S., Sejnowski, T.: Learning overcomplete representation. Neural Comput. **12**(2), 337–365 (2000)
21. Attias, H.: Independent factor analysis. Neural Comput. **11**, 803–851 (1999)
22. Welling, M., Weber, M.: A constrained EM algorithm for independent component analysis. Neural Comput. **13**, 677–689 (2001)
23. Miskin, J.W., MacKay, D.J.C.: Ensemble learning for blind source separation. In: S. Roberts, R. Everson (eds.) Independent Component Analysis: Principles and Practice, pp. 209–233. Cambridge University Press (2001)
24. Bach, F., Jordan, M.I.: Kernel independent component analysis. J. Mach Learn Res. **3**, 1–48 (2002)
25. Plumbley, M.D.: Algorithms for nonnegative independent component anlaysis. IEEE Trans. Neural Netw. **14**(3), 534–543 (2003)
26. Lee, B.I., Lee, J.S., Lee, D.S., Kang, W.J., Lee, J.J., Choi, S.: A clinical application of ensemble ICA to the quantification of myocardial blood flow in dynamic PET. J. VLST Signal Process Syst. Signal Image Video Technol. **49**, 233–241 (2007)
27. Matsuoka, K., Ohya, M., Kawamoto, M.: A neural net for blind separation of nonstationary signals. Neural Netw. **8**(3), 411–419 (1995)
28. Choi, S., Cichocki, A., Amari, S.: Equivariant nonstationary source separation. Neural Netw. **15**(1), 121–130 (2002)
29. Li, Y., Cichocki, A., Amari, S.: Blind estimation of channel parameters and source components for EEG signals: A sparse factorization approach. IEEE Trans. Neural Netw. **17**(2), 419–431 (2006)
30. Hyvärinen, A., Hoyer, P.: Emergence of phase- and shift-invariant features by decomposition of natural images into independent feature subspaces. Neural Comput. **12**(7), 1705–1720 (2000)
31. Bach, F.R., Jordan, M.I.: Beyond independent components: Trees and clusters. J Mach Learn Res. **4**, 1205–1233 (2003)

Independent Factor Analysis

▶ Independent Component Analysis

Indexing

▶ Biometric Algorithms

Individuality of Biometric Traits

▶ Individuality of Fingerprints

Individuality of Fingerprints

Sarat C. Dass[1], S. Pankanti[2], S. Prabhakar[3], Y. Zhu[4]
[1]Michigan State University, East Lansing, MI, USA
[2]IBM T.J. Watson Research Center, Hawthorne, NY, USA
[3]DigitalPersona, Redwood City, CA, USA
[4]Discover Financial Services, Riverwoods, IL, USA

Synonyms

Fingerprint individuality; Fingerprint, Forensic Evidence of; Individuality of biometric traits

Definition

Fingerprint individuality is the study of the extent of uniqueness of fingerprints. It is the most important measure to be ascertained when fingerprint evidence is presented in court by experts. A measure of fingerprint individuality reflects the amount of uncertainty associated with the experts' decision, which arises primarily due to the variability of feature characteristics in a pair of fingerprints. This inherent variability can cause random matching between the pair of fingerprints even if they are not from the same person. Fingerprint individuality aims to characterize this randomness in matching them quantitatively in terms of statistical models.

Introduction

The two fundamental premises on which fingerprint identification is based are: (1) fingerprint details are permanent, and (2) fingerprints of an individual are unique. The validity of the first premise has been established by empirical observations as well as based on the anatomy and morphogenesis of friction ridge skin. It is the second premise that is being challenged in recent court cases. The notion of ▶ fingerprint individuality has been widely accepted based on a manual inspection (by experts) of millions of fingerprints. Based on this notion, expert testimony is delivered in a courtroom by comparing salient features of a latent print lifted from a crime scene with those taken from the defendant. A reasonably high degree of match between the salient features leads the experts to testify irrefutably that the owner of the latent print and the defendant are one and the same person. For decades, the testimony of forensic fingerprint experts was almost never excluded from these cases, and on cross-examination, the foundations and basis of this testimony were rarely questioned. Central to establishing an identity based on fingerprint evidence is the assumption of discernible uniqueness; salient features of fingerprints of different individuals are observably different, and therefore, when two prints share many common features, the experts conclude that the owners of the two different prints are one and the same person. The assumption of discernible uniqueness, although lacking sound theoretical and empirical foundations, allows forensic experts to offer an unquestionable proof toward the defendant's guilt.

A significant event that questioned this trend occurred in 1993 in the case of Daubert v. Merrell Dow Pharmaceuticals [1], where the U.S. Supreme Court ruled that in order for an expert forensic testimony to be allowed in courts, it had to be subject to five main criteria of scientific validation, that is, whether (1) the particular technique or methodology has been subjected to statistical hypothesis testing, (2) its error rates has been established, (3) standards controlling the technique's operation exist and have been maintained, (4) it has been peer reviewed, and (5) it has a general widespread acceptance [2]. Forensic evidence based on fingerprints was first challenged in the 1999 case of USA v. Byron Mitchell [3] under Daubert's ruling, stating that the fundamental premise for asserting the uniqueness of fingerprints had not been objectively tested and its potential matching error rates were unknown. After USA versus Byron Mitchell, fingerprint-based identification has been challenged in more than 20 court cases in the U.S.

The main issue with the admissibility of fingerprint evidence is that the underlying scientific basis of fingerprint individuality has not been rigorously studied or tested. In particular, the central question is: What is the uncertainty associated with the experts' judgment? How likely can an erroneous decision be made for the given latent print? In March 2000, the U.S. Department of Justice admitted that no such testing has been done and acknowledged the need for such a study [4]. In response to this, the National Institute of Justice issued a formal solicitation for "Forensic Friction Ridge (Fingerprint) Examination Validation Studies," whose goal is to conduct "basic research to determine the scientific validity of individuality in friction ridge examination based on measurement of features, quantification, and statistical analysis" [4]. The two main topics of basic research under this solicitation include: (1) the amount of detail in a single fingerprint that is available for comparison, and (2) the amount of detail in correspondence between two fingerprints.

This article gives an overview of the problem of fingerprint individuality, the challenges faced and the models and methods that have been developed to study the extent of uniqueness of a finger. Interest in the fingerprint individuality problem is twofold. First, a scientific basis (a reliable statistical estimate of the matching error) for fingerprint comparison can determine the admissibility of fingerprint

identification in the courts of law as an evidence of identity. Secondly, it can establish an upper bound on the performance of automatic fingerprint verification systems.

The main challenge in assessing fingerprint individuality is to elicit models that can capture the variability of fingerprint features in a population of individuals. Fingerprints are represented by a large number of features, including the overall ridge flow pattern, ridge frequency, location and position of singular points (core(s) and delta(s)), type, direction, and location of minutiae points, ridge counts between pairs of minutiae, and location of pores. These features are also used by forensic experts to establish an identity, and therefore, contribute to the assessment of fingerprint individuality. Developing statistical models on complex feature spaces is difficult albeit necessary. In this chapter, minutiae have been used as the fingerprint feature of choice to keep the problem tractable and as a first step. There are several reasons for this choice: Minutiae is utilized by forensic experts, it has been demonstrated to be relatively stable, and it has been adopted by most of the commonly available automatic fingerprint matching systems. In principal, the assessment of fingerprint individuality can be carried out for any particular matching mode, such as by human experts or by automatic systems, as long as appropriate statistical models are developed on the relevant feature space used in the matching. Thus, the framework also extends to the case where matching is performed based on an automatic system.

Even for the simpler fingerprint feature, namely minutiae, capturing its variability in a population of fingerprints is challenging. For example, it is known that fingerprint minutiae tend to form clusters [5, 6], minutiae information tend to be missed in poor quality images and minutiae location and direction information tend to be highly dependent on one another. All these characteristics of minutiae variability, in turn, affect the chance that two arbitrary fingerprints will match. For example, if the fingerprint pair have minutiae that are clustered in the same region of space, there is a high chance that minutiae in the clustered region will randomly match one another. In this case, the matches are spurious, or false, and statistical models for fingerprint individuality should be able to quantify the likelihood of spurious matches. To summarize, candidate models for assessing fingerprint individuality must meet two important requirements: (1) flexibility, that is, the model can represent the observed distributions of the minutiae features in fingerprint images over different databases, and (2) associated measures of fingerprint individuality can be easily obtained from these models.

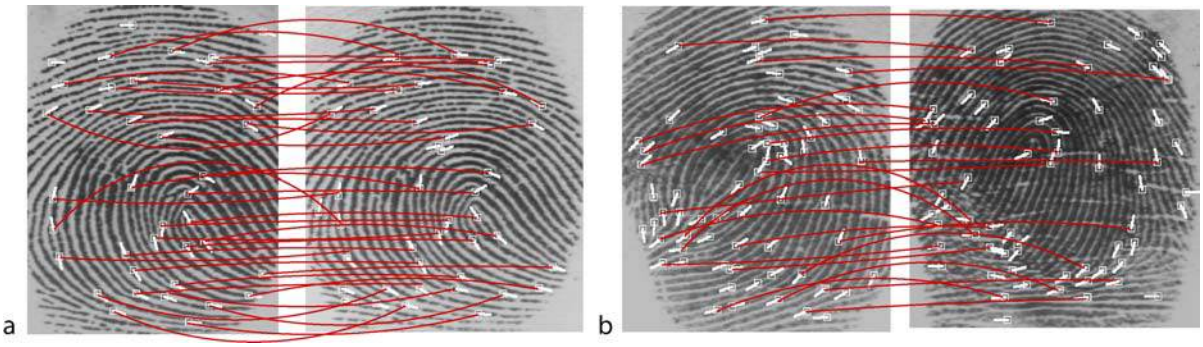
Several works have been reported in the literature on fingerprint individuality. The reader is referred to the overview by Pankanti et al. [2] on this subject. This article focuses on two recent works of fingerprint individuality where statistical models have been developed for minutiae to address the question of fingerprint individuality. These two works are (1) Pankanti et al. [2], and (2) Zhu et al. [7].

The Statistical Test of Biometric Recognition

Fingerprint based recognition, and more generally biometric recognition, can be described in terms of a test of statistical hypotheses. Suppose a query image, Q, corresponding to an unknown identity, I_p, is acquired, fingerprint experts claim that Q belongs to individual I_c, say. This is done by retrieving information of a template image T of I_c and matching T with Q. The two competing expert decision can be stated in terms of two competing hypotheses: The null hypothesis, H_0, states that I_c is not the owner of the fingerprint Q (i.e., Q is an **i**mpostor impression of I_c), and the alternative hypothesis, H_1, states that I_c is the owner of Q (i.e., Q is a *genuine* impression of I_c). The hypotheses testing scenario is

$$H_0 : I_t \neq I_c \quad \text{versus} \quad H_1 : I_t = I_c. \qquad (1)$$

Forensic experts match Q and T based on their degree of similarity (see Fig. 1). For the present scenario, it will be assumed that the degree of similarity is given by the number of matched minutiae pairs, $S(Q,T)$, between Q and T. Large (respectively, small) values of $S(Q,T)$ indicate that T and Q are similar to (respectively, dissimilar to) each other. If $S(Q,T)$ is lower (respectively, higher) than a prespecified threshold λ, it leads to rejection (respectively, acceptance) of H_0. Since noise factors distort information in the prints, two types of errors can be made: False match, which is also called the Type I error in statistics (since H_0 is rejected when it is true.) (FM) and false non-match, also known as the Type II error in statistics (since H_0 is accepted when H_0 is false.) False match



Individuality of Fingerprints. Figure 1 Illustrating genuine and impostor minutiae matching (taken from [2]). (**a**) Two impressions of the same finger are matched; 39 minutiae were detected in input (*left*), 42 in template (*right*), and 36 "true" correspondences were found. (**b**) Two different fingers are matched; 64 minutiae were detected in input (*left*), 65 in template (*right*), and 25 "false" correspondences were found. © 2002 IEEE.

occurs when an expert incorrectly accepts an impostor print as a match whereas false nonmatch occurs when the expert incorrectly rejects a genuine fingerprint as a nonmatch. The false match and nonmatch rates (FMR and FNMR, respectively), are the probability of FM and FNM. The formulae for FMR and FNMR are:

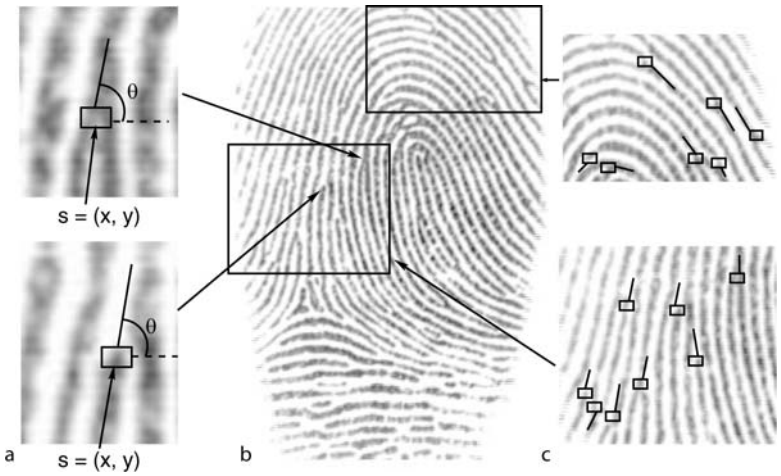
$$\begin{aligned} \mathrm{FMR}(\lambda) &= P(S(Q,T) > \lambda \,|\, I_t \neq I_c), \\ \mathrm{FNMR}(\lambda) &= P(S(Q,T) \leq \lambda \,|\, I_t = I_c). \end{aligned} \tag{2}$$

In case there is no external noise factors that affect the acquisition of Q and T, it can decided without error whether Q belongs to I_c or not based on the premise of the uniqueness of fingerprints. However, the process of fingerprint acquisition is prone to many sources of external noise factors that distort the true information present in Q (as well as T). For example, there can be variability due to the placement of the finger on the sensing plane, smudges and partial prints in the latent that is lifted from the crime scene, nonlinear distortion due to the finger skin elasticity, poor quality image due to dryness of the skin, and many other factors. These noise factors cause information in Q to be distorted, for example, true minutiae points may be missed and spurious minutiae points can be generated, which in turn affects the uncertainty associated with rejecting or accepting H_0.

The different noise factors can be grouped into two major sources of variability: (1) inter- and (2) intra-class fingerprint variability. Intraclass variability refers to the fact that fingerprints from the same finger look different from one another. As mentioned earlier, sources for this variability includes nonlinear deformation due to skin elasticity, partial print, non-uniform fingertip pressure, poor finger-condition (e.g., dry finger), noisy environment, etc. Interclass variability refers to the fact that fingerprints from different individuals look very similar. Unlike intra-class variability, the cause of interclass variability is intrinsic to the target population. Panels (b) of Fig. 1 show an example of interclass variability for two different fingerprint images. Both intra- and interclass variability need to be accounted for, when determining whether Q and T match or not. It is easy to see that fingerprint experts will be able to make more reliable decisions if the inter-class fingerprint variability is large and the intra-class fingerprint variability is small. However, less reliable decisions will be made if the reverse happens, that is, when intraclass variability is large and interclass variability is small. In other words, the study of fingerprint individuality is the study of quantification of inter- and intraclass variability in Q and T, as well as to what extent these sources of variability affect the fingerprint expert's decision.

Statistical Models for Fingerprint Individuality

The study and quantification of inter- and intraclass variability can be done by eliciting appropriate stochastic (or, statistical) models on fingerprint minutiae. Figure 2 shows two examples of minutiae (ending and bifurcation) and the corresponding location and direction information. Two such approaches are described in this section, namely, the work done by Pankanti et al. [2] and the subsequent model that was proposed by Zhu et al. [7]. Both works focus on modeling the



Individuality of Fingerprints. **Figure 2** Minutiae features consisting of the location, s, and direction, θ, for a typical fingerprint image (**b**): The top (respectively, bottom) panel in (**a**) shows s and θ for a ridge bifurcation (respectively, ending). The top (respectively, bottom) panel in (**a**) shows two subregions in which orientations of minutiae points that are spatially close tend to be very similar. © 2007 IEEE.

I

interclass fingerprint variability, that is, the variability inherent in fingerprint minutiae of different fingers in a population.

Pankanti's Fingerprint Individuality Model

The set up of Pankanti et al. [2] is as follows: Suppose the query fingerprint Q has n minutiae and the template T has m minutiae denoted by the sets

$$M_Q \equiv \{\{S_1^Q, D_1^Q\}, \{S_2^Q, D_2^Q\}, ..., \{S_n^Q, D_n^Q\}\} \quad (3)$$

$$M_T \equiv \{\{S_1^T, D_1^T\}, \{S_2^T, D_2^T\}, ..., \{S_m^T, D_m^T\}\}, \quad (4)$$

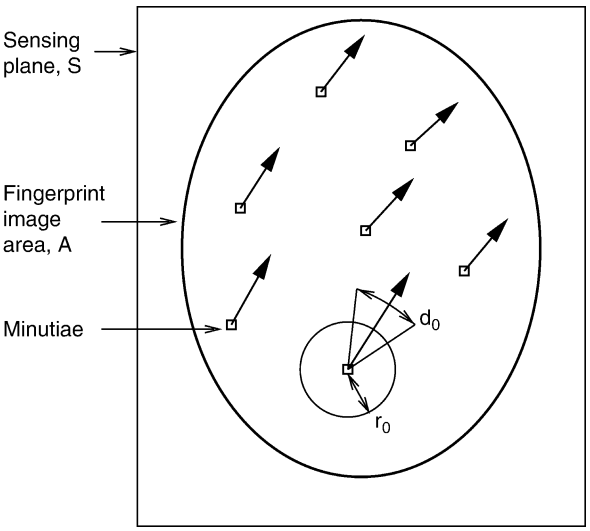
where in (3) and (4), S and D refer to a generic minutiae location and direction pair. To assess a measure of fingerprint individuality, it is first necessary to define a minutiae correspondence between Q and T. A minutiae in Q, (S^Q, D^Q), is said to match (or, correspond) to a minutiae in T, (S^T, D^T), if for fixed positive numbers r_0 and d_0, the following inequalities are valid:

$$|S^Q - S^T|_s \le r_0 \quad \text{and} \quad |D^Q - D^T|_d \le d_0, \quad (5)$$

where

$$|S^Q - S^T|_s \equiv \sqrt{(x^Q - x^T)^2 + (y^Q - y^T)} \quad (6)$$

is the Euclidean distance between the minutiae locations $S^Q \equiv (x^Q, y^Q)$ and $S^T \equiv (x^T, y^T)$, and



Individuality of Fingerprints. **Figure 3** Identifying the matching region for a query minutiae (image taken from [2] and [7]). © 2002 IEEE.

$$|D^Q - D^T|_d \equiv \min(|D^Q - D^T|, 2\pi - |D^Q - D^T|) \quad (7)$$

is the angular distance between the minutiae directions D^Q and D^T. The choice of parameters r_0 and d_0 defines a tolerance region (see Fig. 3), which is critical in determining a match according to (5). Large (respectively, small) values of the pair (r_0, d_0) will lead to

spurious (missed) minutiae matches. Thus, it is necessary to select (r_0, d_0) judiciously so that both kinds of matching errors are minimized. A discussion on how to select (r_0, d_0) is given subsequently.

In [2], fingerprint individuality was measured in terms of the probability of random correspondence (PRC). The PRC of w matches is the probability that two arbitrary fingerprints from a target population have at least w pairs of minutiae correpondences between them. Recall the hypothesis testing scenario of (1) for biometric authentication. When the similarity measure $S(Q,T)$ is above the threshold λ, the claimed identity (I_c) is accepted as true identity. Based on the statistical hypothesis in (1), the PRC is actually the false match rate, FMR, given by

$$\mathrm{PRC}(w) = P(S(Q,T) \geq w \,|\, I_c \neq I_t) \tag{8}$$

evaluated at $\lambda = w$.

To estimate the PRC, the following assumptions were made in [2]: (1) Only minutiae ending and bifurcation are considered as salient fingerprint features for matching. Other types of minutiae, such as islands, spur, crossover, lake, etc., rarely appear and can be thought of as combination of endings and bifurcations. (2) Minutiae location and direction are uniformly distributed and independent of each other. Further, minutiae locations cannot occur very close to each other. (3) Different minutiae correspondences between Q and T are independent of each other, and any two correspondences are equally important. (4) All minutiae are assumed true, that is there are no missed or spurious minutiae. (5) Ridge width is unchanged across the whole fingerprint. (6) Alignment between Q and T exists, and can be uniquely determined.

Based on the above assumptions, Pankanti et al. were able to come up with the uniform distribution as the statistical model for fingerprint individuality. The probability of matching w minutiae in both position as well as direction is given by

$$p(M,m,n,w) = \sum_{\rho=w}^{\min(m,n)} \left(\frac{\binom{m}{\rho}\binom{M-m}{n-\rho}}{\binom{M}{n}} \times \binom{\rho}{w} (l)^w (1-l)^{\rho-w} \right), \tag{9}$$

where $M = A/C$ with A and C defined, respectively, as the area of overlap between Q and T and $C = \pi r_0^2$ is the area of the circle with radius r_0. Pankanti et al. further improved their model based on several considerations of the occurrence of minutiae. The ridges occupy approximately $\frac{A}{2}$ of the total area with the other half occupied by the valleys. Assuming that the number (or the area) of ridges across all fingerprint types is the same and that the minutiae can lie only on ridges, i.e., along a curve of length $\frac{A}{\omega}$, where ω is the ridge period, the value of M in (9) is changed from $M = A/C$ to

$$M = \frac{A/\omega}{2r_0}, \tag{10}$$

where $2r_0$ is the length tolerance in minutiae location.

Parameters (r_0, d_0) determine the minutiae matching region. In the ideal situation, a genuine pair of matching minutiae in the query and template will correspond exactly, which leads to the choice of (r_0, d_0) as (0,0). However, intraclass variability factors such as skin elasticity and nonuniform fingertip pressure can cause the minutiae pair that is supposed to perfectly match, to slightly deviate from one another. To avoid rejecting such pairs as nonmatches, nonzero values of r_0 and d_0 need to be specified for matching pairs of genuine minutiae. The value of r_0 is determined based on the distribution of the Euclidean distance between every pair of matched minutiae in the genuine case. To find the corresponding pairs of minutiae, pairs of genuine fingerprints were aligned, and Euclidean distance between each of the genuine minutiae pairs was then calculated. The value of r_0 was selected so that only the upper 5% of the genuine matching distances (corresponding to large values of r) were rejected. In a similar fashion, the value of d_0 was determined to be the 95th percentile of this distribution (i.e., the upper 5% of the ▶ genuine matching angular distances were rejected).

To find the actual r_0 and d_0, Pankanti et al. used a database of 450 mated fingerprint pairs from IBM ground truth database (see [2] for details). The true minutiae locations in this database and the minutiae correspondences between each pair of genuine fingerprints in the database were determined by a fingerprint expert. Using the ground truth correspondences, r_0 and d_0 were estimated to be 15 and 22.5, respectively. These values will be used to estimate the PRC in the experiments presented in this paper.

Pankanti et al. [2] were the first to attempt in quantifying a measure of fingerprint individuality

based on statistical models. However, the proposed uniform model does have some drawbacks. Comparison between model prediction and empirical observations showed that the corrected uniform model grossly underestimated the matching probabilities. The inherent drawbacks of the uniform model motivated the research by Zhu et al. [7] to propose statistical distributions that can better represent minutiae variability in fingerprints.

Mixture Models for Fingerprint Features

Zhu et al. [7] proposed a mixture model to model the minutiae variability of a finger by improving Assumption (2) of [2]. A joint distribution model for the k pairs of minutiae features $\{(S_j, D_j), j = 1, 2, \ldots k\}$ is proposed to account for (1) clustering tendencies (i.e., nonuniformity) of minutiae, and (2) dependence between minutiae location (S_j) and direction (D_j) in different regions of the fingerprint. The mixture model on (S, D) is given by

$$f(s,\theta\,|\,\Theta_G) = \sum_{g=1}^{G} \tau_g f_g^S(s\,|\,\mu_g, \Sigma_g) \cdot f_g^D(\theta\,|\,v_g, \kappa_g), \quad (11)$$

where G is the total number of mixture components, $f_g^S(\cdot)$ is the bivariate Gaussian density with mean μ_g and covariance matrix Σ_g, and

$$f_g^D(\theta\,|\,v_g, \kappa_g, p_g) = \begin{cases} p_g\, \nu(\theta) & \text{if } 0 \le \theta < \pi \\ (1-p_g)\, \nu(\theta-\pi) & \text{if } \pi \le \theta < 2\pi, \end{cases} \quad (12)$$

where $\nu(\theta)$ is the Von-Mises distribution for the minutiae direction given by

$$\nu(\theta) \equiv \nu(\theta\,|\,v_g, \kappa_g) = \frac{2}{I_0(\kappa_g)} \exp\{\kappa_g \cos 2(\theta - v_g)\} \quad (13)$$

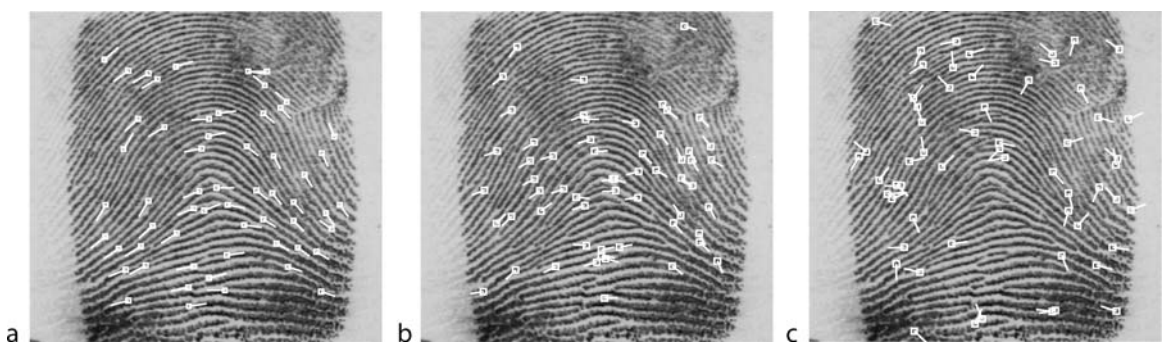
with $I_0(\kappa_g)$ defined as

$$I_0(\kappa_g) = \int_0^{2\pi} \exp\{\kappa_g \cos(\theta - v_g)\}\, d\theta. \quad (14)$$

In (13), v_g and κ_g represent the mean angle and the precision (inverse of the variance) of the Von-Mises distribution, respectively (see [7] for details). The distribution f_g^D in (12) can be interpreted in the following way: The ridge flow orientation, o, is assumed to follow the Von-Mises distribution in (13) with mean v_g and precision κ_g. Subsequently, minutiae arising from the gth component have directions that are either o or $o+\pi$ with probabilities p_g and $1-p_g$, respectively.

The model described by (11) has three distinct advantages over the uniform model: (1) it allows for different clustering tendencies in minutiae locations and directions via G different clusters, (2) it incorporates dependence between minutiae location and direction since, if S is known to come from the gth component, the direction D also comes from the gth component, and (3) it is flexible in that it can fit a variety of observed minutiae distributions adequately. The estimation of the unknown parameters in (11) has been described in details in [7].

The effectiveness of the mixture models can also be shown by simulating from the fitted models and checking to see if a similar pattern of minutiae is obtained as observed. Figure 4a shows a fingerprint whose



Individuality of Fingerprints. Figure 4 All (S,D) realizations from the fitted mixture model (**b**), and from the uniform distribution (**c**) for the original image in (**a**). The true minutiae locations and directions are marked in (**a**). Images are taken from [7]. © 2007 IEEE.

minutiae features were fitted with the mixture distribution in (11). Figure 4c shows a simulated realization when each S and D is assumed to be uniformly distributed independent of each other. Note that there is a good agreement, in the distributional sense, between the observed (Fig. 4a) and simulated minutiae locations and directions from the fitted mixture model (Fig. 4b), but no such agreement exists for the uniform model.

Zhu et al. [7] obtains a closed form expression for the PRC corresponding to w matches under similar assumptions of Pankanti et al. [2] (barring Assumption (2)). The probability of obtaining exactly w matches given there are m and n minutiae in Q and T, respectively, is given by the expression

$$p^*(w;\, Q, T) = \frac{e^{-\lambda(Q,T)}\,\lambda(Q,T)^w}{w!} \tag{15}$$

for large m and n; (15) corresponds to the Poisson probability mass function with mean $\lambda(Q, T)$ given by

$$\lambda(Q, T) = m\, n\, p(Q, T), \tag{16}$$

where

$$p(Q, T) = P(|S^Q - S^T|_s \le r_0 \text{ and } |D^Q - D^T|_a \le d_0) \tag{17}$$

denotes the probability of a match when (S^Q, D^Q) and (S^T, D^T) are random minutiae from the mixture distributions fitted to Q and T, respectively. The mean parameter $\lambda(Q, T)$ can be interpreted as the expected number of matches from the total number of mn possible pairings between m minutiae in Q and n minutiae points in T with the probability of each match being $p(Q, T)$.

Incorporating Interclass Variability Via Clustering

The above PRC was obtained for a single query and template fingerprint pair. An important difference between the proposed methodology and previous work is that mixture models are fitted to each finger, whereas previous studies assumed a common distribution for all fingers/impressions. Assuming a common minutiae distribution for all fingerprint impressions has a serious drawback, namely, that the true distribution of minutiae may not be modeled well and important cluster information may be smoothed out. Zhu et al. [7] adopt an agglomerative hierarchical clustering procedure on the space of all fitted mixture models to obtain a reliable representation of the minutiae in all fingerprints in a population as well as to reduce computational time to obtain the PRC estimates. In this framework, the probability of obtaining exactly u matches corresponding to clusters C_i and C_j is given by

$$p^*(u;\, C_i, C_j) = e^{-\lambda(C_i,C_j)}\frac{\lambda(C_i, C_j)^u}{u!}, \tag{18}$$

where $\lambda(C_i, C_j)$ is interpreted as the mean number of matches between any Q arising from C_i and T arising from C_j, and is calculated based on the mean (average) mixture densities in clusters C_i and C_j. The overall probability of exactly u matches is

$$p^{**}(u) = \frac{\sum_{i\le j} |C_i|\,|C_j|\, p^*(u;\, C_i, C_j)}{\sum_{i\le j} |C_i|\,|C_j|,} \tag{19}$$

where $|C_k|$ is the total number of mixtures in cluster C_k. The overall PRC corresponding to w matches is given by

$$\overline{\text{PRC}} = \sum_{u=w}^{\infty} p^{**}(u). \tag{20}$$

To remove the effect of very high or very low PRCs, the $100(1-\alpha)\%$ trimmed mean is used instead of the ordinary mean as in (19). The lower and upper $100\alpha/2$th percentiles of $\{p^*(u; C_i, C_j), 1 \le i, j \le N^*\}$ are denoted by $p_C{}^*(u;\alpha/2)$ and $p_C{}^*(u;1-\alpha/2)$. Also, define the set of all trimmed $p^*(u; C_i, C_j)$ probabilities as $\mathcal{T} \equiv \{(i,j) : p_C^*(u;\alpha/2) \le p^*(u;\, C_i, C_j) \le p_C^*(u;1-\alpha/2)\}$. Then, the $100(1-\alpha)\%$ trimmed mean PRC is

$$\overline{\text{PRC}}_\alpha = \sum_{u=w}^{\infty} p_{\mathcal{T}}^{**}(u), \tag{21}$$

where

$$p_{\mathcal{T}}^{**}(u) = \frac{\sum_{(i,j)\in\mathcal{T}} |C_i|\,|C_j|\, p^*(u;\, C_i, C_j)}{\sum_{(i,j)\in\mathcal{T}} |C_i|\,|C_j|\,.} \tag{22}$$

In the next section, the trimmed mean is with $\alpha = 0.05$.

Experimental Results

The results in this section are taken from Zhu et al. [7]; the interested reader is referred to more details

discussed in the chapter. The methodology for assessing the individuality of fingerprints are validated on three target populations, namely, the NIST Special Database 4 [8], FVC2002 DB1 and FVC2002 DB2 [9] fingerprint databases. The NIST fingerprint database [8] is publicly available and contains 2,000 8-bit gray scale fingerprint image pairs of size 512 × 512 pixels. Because of the relative large size of the images in the NIST database, the first image of each pair is used for statistical modeling. Minutiae could not be automatically extracted from two images of the NIST database due to poor quality. Thus, the total number of NIST fingerprints used in the authors' experiments is $F = 1{,}998$.

For the FVC2002 database, also available in the public domain, two of its subsets DB1 and DB2 is used. The DB1 impressions (images size = 388 × 374) are acquired using the optical sensor "TouchView II" by Identix, while the DB2 impressions (image size = 296 × 560) are acquired using the optical sensor "FX2000" by Biometrika. Each database consists of $F = 100$ different fingers with eight impressions ($L = 8$) per finger. Because of the small size of the DB1 and DB2 databases, a minutiae consolidation procedure was adopted to obtain a master (see [7] for the details). The mixture models were subsequently fitted to each master.

Zhu et al. developed a measure of goodness of fit of hypothesized distributions to the observed minutiae based on a chi-square type criteria. Two tests were considered, namely, the Freeman–Tukey and Chi-square tests. The results for the goodness of fit for two hypothesized distributions, namely, mixture and uniform models are reported in Table 1. For all the three databases, the number of fingerprint images with *p*-values above (corresponding to accepting the hypothesized distribution) and below the threshold 0.01 (corresponding to rejecting the hypothesized distribution) were obtained. Note that the entries in Table 1 imply that the mixture model is generally a better fit to the observed minutiae compared to the uniform; for example, the mixture model is a good fit to 1,666 images from the NIST database (corresponding to *p*-values above 0.01) based on the Freeman–Tukey test. For the Chi-square test, this number is 1,784. In comparison, the uniform model is a good fit to only 905 and 762 images, respectively.

Zhu et al. compared the PRC obtained by [7] with those of Pankanti et al. [2]. The query and template fingerprints in the NIST and FVC databases are first aligned using the matcher described in [10], and an overlapping area between the two fingerprints are determined. To compute the PRCs, the mixture models are restricted onto overlapping area (see [7] for more details). Table 2 gives the mean m, mean n, mean overlapping area and M (see equation (10)) for the NIST and Fvc databses, whereas Table 3 gives the corresponding PRCs. The empirical PRC is computed as the proportion of impostor pairs with 12 or greater matches among all pairs with m and n values within ± 5 of the mean in the overlapping area. The empirical probabilities of at least w matches are obtained by counting the number of fingerprint pairs with 12 or more matches divided by the total number of pairs. Thus, one should note that the empirical probability is

Individuality of Fingerprints. **Table 1** Results from the Freeman–Tukey and Chi-square tests for testing the goodness of fit of the mixture and uniform models

Mixture model						
	Freeman–Tukey			Chi-square		
p-value	NIST (1,998)	DB1 (100)	DB2 (100)	NIST (1,998)	DB1 (100)	DB2 (100)
p-value > 0.01 (mixture accepted)	1,864	71	67	1,569	65	52
p-value ≤ 0.01 (mixture rejected)	134	29	33	429	35	48
Uniform model						
	Freeman–Tukey			Chi-square		
p-value	NIST (1,998)	DB1 (100)	DB2 (100)	NIST (1,998)	DB1 (100)	DB2 (100)
p-value > 0.01 (uniform accepted)	550	1	0	309	1	0
p-value ≤ 0.01 (uniform rejected)	1,448	99	100	1,689	99	100

Entries correspond to the number of fingerprints in each database with *p*-values above and below 0.01. The total number of fingerprints in each database is indicated in parenthesis. Table entries are taken from [7]

Individuality of Fingerprints. **Table 2** Table giving the mean m and n in the overlapping area, the mean overlapping area and the value of M for each database [7]

Database	(m,n)	Mean overlapping area (pixel^2)	M
NIST	(52,52)	112,840	413
FVC2002 DB1	(51,51)	71,000	259
FVC2002 DB2	(63,63)	110,470	405

Individuality of Fingerprints. **Table 3** A comparison between fingerprint individuality estimates using the (a) Poisson and mixture models, and (b) Pankanti et al. [2]

Database	(m,n,w)	Empirical		Mixture		Pankanti	
		Mean no. of matches	PRC	Mean λ	PRC	Mean λ	PRC
NIST	(52,52,12)	7.1	3.9×10^{-3}	3.1	4.4×10^{-3}	1.2	4.3×10^{-8}
FVC2002 DB1	(51,51,12)	8.0	2.9×10^{-2}	4.9	1.1×10^{-2}	2.4	4.1×10^{-6}
FVC2002 DB2	(63,63,12)	8.6	6.5×10^{-2}	5.9	1.1×10^{-2}	2.5	4.3×10^{-6}

matcher dependent. Since fingerprint individuality is assessed based on minutiae location and direction only, the matcher of [10] was used which depends only on minutiae information.

Note that as m or n or both increase, the values of PRCs for both the models become large as it becomes much easier to obtain spurious matches for larger m and n values. Additionally, Table 3 illustrates an important fact: The PRCs based on the mixture models are orders of magnitude larger compared to Pankanti's model and closer to the empirical probability of at least w matches. Note also that the mean of λs (the theoretical mean number of matches) are closer to the empirical counterpart (mean number of observed matches) compared to Pankanti's model. This demonstrates the adequateness of the mixture models for the assessment of fingerprint individuality. While the mixture models is more adequate at representing minutiae variability, the PRCs obtained are far too large indicating a large amount of uncertainty in declaring a match between a fingerprint pair. One way to reduce the PRC is to add more fingerprint features when performing the identification. Fingerprint individuality assessment can then be made by developing appropriate statistical models for these features.

Summary and Future Work

In this chapter, an overview of the challenges involved in assessing the individuality of fingerprints is presented. Two works have been discussed. Pankanti's model is the first attempt at modeling the observed minutiae distribution via statistical models, whereas Zhu et al. developed more flexible models that adequately describe all minutiae characteristics. There are many open problems that still remain unsolved. Both works have only addressed the issue of interclass minutiae variability. Appropriate statistical models for modeling the intraclass minutiae variability are still very few in the literature. A very important source of intraclass variability is the quality of the query and template images, and work still needs to be done to investigate how PRCs change with quality of the fingerprint image. It is also important to develop statistical models for more complex fingerprint features. In this case, one can use more useful matching criteria that utilize richer fingerprint features.

In this chapter, the PRCs have been related to the false match rates. Another measure of fingerprint individuality should be related to the false nonmatch rates. Eventually, a measure of fingerprint individuality should be a optimal combination of the two measures of errors.

Acknowledgments

The authors like to thank Prof. Anil Jain for introduction of the fingerprint individuality problem to them and for many subsequent discussions that has helped them in their research in this area. This chapter was written under the support of the NSF DMS grant 0706385.

Related Entries

- ▶ Fingerprint
- ▶ Fingerprint Individuality
- ▶ Fingerprint Matching, Automatic
- ▶ Fingerprint Matching, Manual
- ▶ Forensic Evidence of Individuality

References

1. Daubert v. Merrel Dow Pharmaceuticals Inc.: 509 U.S. 579, 113 S. Ct. 2786, 125, L.Ed.2d 469 (1993)
2. Pankanti, S., Prabhakar, S., Jain, A.K.: On the individuality of fingerprints. IEEE Trans. Pattern Anal. Mach. Intell. **24**(8), 1010–1025 (2002)
3. United States v. Byron Mitchell: Criminal Action No. 96-407, United States District Court for the Eastern District of Pennsylvania (1999)
4. United States Department of Justice: Document SL000386, March 2000. Online: http://www.forensic-evidence.com/site/ID/IDfpValidation.html
5. Scolve, S.C.: The occurence of fingerprint characteristics as a two dimensional process. J. Am. Stat. Assoc. **74**(367), 588–595 (1979)
6. Stoney, D.A., Thornton, J.I.: A critical analysis of quantitative fingerprint individuality models. J. Forensic Sci. **31**(4), 1187–1216 (1986)
7. Zhu, Y., Dass, S.C., Jain, A.K.: Statistical models for assessing the individuality of fingerprints. IEEE Trans. Inf. Forensics Secur. **2(3)**, 391–401 (2007)
8. NIST: 8-bit gray scale images of fingerprint image groups (FIGS). Online: http://www.nist.gov/srd/nistsd4.htm
9. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2002: Fingerprint verification competition. In: Proceedings of the International Conference on Pattern Recognition (ICPR), pp. 744–747 (2002)
10. Ross, A., Dass, S., Jain, A.K.: A deformable model for fingerprint matching. Pattern Recognit **38**(1), 95–103 (2005)

Individualization

▶ Fingerprint Matching, Manual

Influencing Factors

▶ Influential Factors to Performance

Influential Factors to Performance

Kaoru Uchida
NEC Corporation, Kawasaki, Japan

Synonym

Influencing factors

Definition

Factors that influence biometric performance. They can be discussed from the following viewpoints: characteristics of users (including the definition of impostors) and restrictions that come from practical situations in which that biometric modality is used in applications.

Introduction

When evaluating performance of biometric systems, factors that influence performance ("influencing factors") should be identified and analyzed because performance is greatly affected by a wide variety of influencing factors. The same biometric device may generate different test results if these influencing factors differ. Controlling, recording, and reporting factors are indispensable for executing repeatable performance tests and predicting operational performance [1–3].

Influencing Factors

Identifying influencing factors: the following factors should be considered at a minimum (see also Annex C of [1]):

1. Biometric sensor quality and characteristics
2. Biological or behavioral characteristics of the subject relevant to data collection (essential historical or demographic data):
   a. Invariable: Gender, ethnic origin
   b. Variable:
      - Biological: age, body dimensions/anthropometric data (height, weight, etc. . .), musculoskeletal disorders

I

- Habitual/Social factors: smoking preference, hairstyle, makeup, eyewear (glasses, contacts, etc. . .), clothing
- Occupation

3. Environmental factors applicable to the biometric device, sensor, or application such as:
   a. Temperature
   b. Humidity
   c. Illumination
      - Type (standard incandescent, fluorescent, tungsten halogen, reflector lamps, light emitting diodes (LEDs), sunlight, etc. . .)
   d. Noise
   e. Position of sensor with regards to the user
4. Temporal change of the biometric features
5. Impact of ► forgery attempts on false acceptance, particularly in behavioral modalities
6. Differences between the data capture and signal processing subsystems used in the enrolment phase and those used in the verification/identification phase. (This text taken ISO/IEC TR 19795-3:2007 Information technology – Biometric performance testing and reporting – Part 3: Modality-specific testing is reproduced with the permission of the International Organization for Standardization, ISO. This standard can be obtained from any ISO member and from the Web site of the ISO Central Secretariat at the following address: www.iso.org. Copyright remains with ISO).

In addition, the characteristics that affect performance can be discussed from the following aspects:

1. The definition of impostors
2. Restrictions that come from practical situations in which that biometric modality is used in applications

Characteristics of Impostors

There are two factors to consider the definition of impostors:

1. Multiple biometric data from one person
2. Impostor attempts for behavior-based modalities, such as voice or signature

For modalities in which multiple biometric data can be collected from one person, e.g., finger (10 fingerprints from one person) and iris (2 iris-images from one person), a rule for permitting or prohibiting the use of these data as impostor attempts should be clearly defined in performance testing.

In the case of behavior-based modalities, testing results regarding impostor attempts (FMR or FAR) may be influenced depending on whether (or how much) an impostor tries to imitate an authorized user's behavior. For instance, the case in which an impostor physically traces an authorized user's signature that the impostor obtained differs significantly in FMR or FAR from the case, where the impostor only looks at the signature and imitates it. For these modalities, a criterion regarding impostor attempts should be defined in performance testing.

Characteristics of Modality Specific to Applications

In general, almost all modalities of biometrics are used for user authentication, but some modalities are expected to be used in different classes of applications, for example, face-based identification is widely used in surveillance applications. While a user's cooperation can be expected in the former, it cannot be expected in the latter case. Thus, variation of performance testing should be considered depending on the way the modality is used in real applications.

These restrictions can be divided into two classifications:

1. Factors relating to users, such as facial expressions that affect the countenance of the face, wearing eye-glasses or contact lenses for the iris, etc.
2. Factors relating to external environments that are uncontrollable by the algorithm or system, such as illumination change for face or background noise for voice

These factors naturally affect the performance, and the types and number of factors are different in each modality. These modality-dependent variations should be considered in performance testing. In addition, a concept of ► robustness test should be introduced to evaluate the sensitivity or robustness of the technology toward environmental factors, in case the variation of the factors strongly influences the observed performance.

Related Entries

- ▶ Performance Evaluation, Overview
- ▶ Performance Measures
- ▶ Performance Testing Methodology, Standardization

References

1. ISO/IEC 19795-1:2006 Information technology – Biometric performance testing and reporting – Part 1: Principles and framework
2. ISO/IEC 19795-2:2007 Information technology – Biometric performance testing and reporting – Part 2: Testing methodologies for technology and scenario evaluation
3. ISO/IEC TR 19795-3:2007 Information technology – Biometric performance testing and reporting – Part 3: Modality-specific testing

Information Content of Iris Images

- ▶ Iris Image Quality

Information Fusion

Information fusion is the merging of disparate information, usually from different sources, which is combined in order to achieve a wider perspective on a problem than is possible if only one type of information were considered. In the context of multimodal biometric authentication, the term is used to describe how information from different biometric modalities, for example a facial image and a voice sample, can be combined to make a better decision on whether the person seeking to be authenticated is who he or she claims to be. Depending on where in the system the information from the two modalities is combined, one can distinguish between *feature fusion*, *score fusion,* and *decision fusion*. Feature fusion is also known as *early fusion*, while the latter two fusion types are known collectively as *late fusion*. Feature fusion is a method, by which the features from the different modalities are fused into combined feature vectors after the feature extraction stage, but before the pattern recognition stage of the authentication system. The pattern recognition process then proceeds on the basis of the combined feature vectors. Score fusion is a method, by which a separate authentication system for each modality calculates a mathematical likelihood or distance score between the feature and the client model or template in that modality. The score fusion algorithm then combines the two modality scores by means of a suitable formula, such as the arithmetic mean of the two scores, and the final accept–reject decision is made on the basis of the combined score. Decision fusion is a method, by which each modality has its own independent authentication process – from feature extraction through score calculation to final accept–reject decision for that modality. The overall accept–reject decision by the system is then made through a simple logical combination of the two modality decisions. Two decision fusion paradigms are possible in a bimodal system: a strict system requires an "accept" from both modalities for an overall accept decision, while a generous system only requires an "accept" from either modality for an overall accept decision. A system, which is to detect synchrony between the audio and video signals from a speaking face, requires early fusion, while late fusion is sometimes used when two off-the-shelf commercial authentication systems are combined and the user has no access to the internal feature vectors of the two systems.

- ▶ Liveness Assurance in Face Authentication

Intelligent Agents

Software (sub-)systems that may act autonomously on behalf of users of the system.

- ▶ Biometric Systems, Agent-Based

Interaction

The result an individual, using a biometric sensor, creates by presenting biometric characteristics to a

sensor. In terms of the general biometric model, this occurs in the data capture silo. During an individual's interaction, the biometric sensor or device acquires an image or signal of biometric characteristics of a person. The goal of the presentation is for an individual to present the biometric characteristics of high quality to the sensor in a repeatable and consistent manner, so that the subsequent signal processing sub-processes of segmentation, feature extraction, and quality control can successfully occur. See also attempt, presentation, transaction, and general biometric model.

▶ Ergonomic Design for Biometric Systems

Interactive Voice Response (IVR)

Interactive voice response (IVR) is an automated, "self service" telephony technology that interacts with a caller to gather and dispense information, perform transactions, or route the call to the appropriate recipient. IVR systems can accept touchtone (DTMF), spoken (speech recognition) input, or a combination of the two modalities.

▶ Remote Authentication
▶ Speaker Recognition, Standardization

Interest Point, Region, Local Feature

In a way, the ideal local feature is a point as defined in geometry: having a location in space but no spatial extent. In practice however, images are discrete with the smallest spatial unit being a pixel and discretization effects playing an important role. To localize features in images, a local neighborhood of pixels need to be analyzed, giving all local features some implicit spatial extent. For some applications (e.g., camera calibration or 3D reconstruction), this spatial extent is completely ignored in further processing, and only the location derived from the feature extraction process is used. In those cases, one typically uses the term interest point. However, in most applications those features also need to be described, such that they can be identified and matched, and this again calls for a local neighborhood of pixels. Often, this neighborhood is taken equal to the neighborhood used to localize the feature, but this need not be the case. In this context, one typically uses the term region instead of interest point. However, beware: when a local neighborhood of pixels is used to describe an interest point, the feature extraction process has to determine not only the location of the interest point, but also the size and possibly the shape of this local neighborhood. Especially in case of geometric deformations, this significantly complicates the process, as the size and shape have to be determined in an invariant (covariant) way.

▶ Local Image Features

Interest Points

▶ Local Image Features

Interface

▶ User Interface, System Design

Intermediate Biometrics

▶ Biometric Sample Synthesis

Internal Identification

Identification of a victim based on medical findings by an autopsy, such as implants, and missing organs.

▶ Dental Biometrics

International Association for Identification

▶ Fingerprint, Forensic Evidence of

Interoperability

Interoperability in biometrics can be defined as the capability of a recognition system to operate with data from different sources (e.g., data acquired using different sensors or features extracted using systems from different vendors). Most biometric systems are designed under the assumption that the data to be compared with are obtained from a unique source and are restricted in their ability to match or compare with biometric data originated from different sources. As a result, changing the source may affect the performance of the system.

▶ Fingerprint Databases and Evaluation
▶ Iris Device

Interoperable Performance

Patrick Grother
National Institute of Standards and Technology, MD, USA

Definition

Accuracy of a biometric system that includes standardized components from several suppliers.

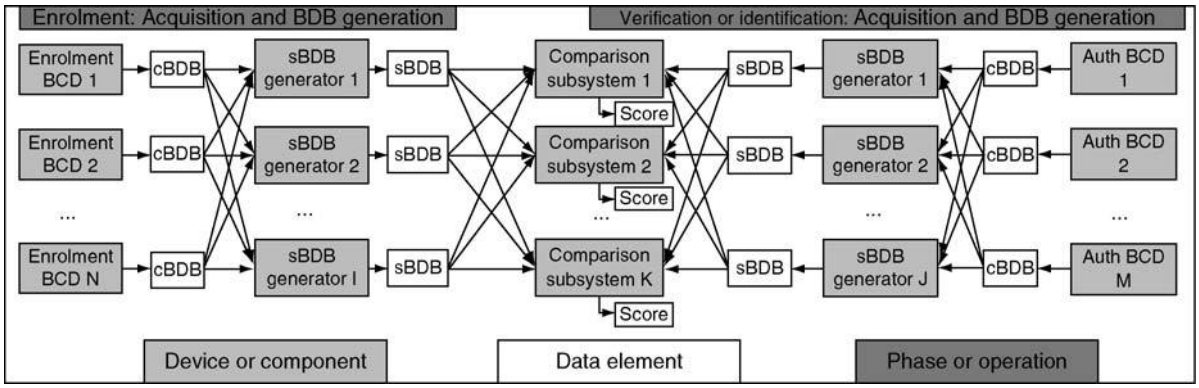
In applications where components conform to standardized interfaces and functional specifications, it is possible to replace one component with another from a different manufacturer. Although conformity to specifications is a necessary condition for interoperability, it is often not sufficient, because the internal algorithmic action of the component is usually not regulated by the standard. Thus, a biometric detection algorithm might underperform some others despite being in conformity with the requirements. This article suggests that the appropriate means of quantifying biometric interoperability are to identify relevant performance metrics, to measure them, and to certify against them. For biometric sensors and also for detection, segmentation, and matching algorithms, these metrics will be usually be failure-to-acquire and enroll rates, and Type I and II recognition error rates.

Introduction

Biometric recognition is explicitly a two-phase operation: In verification, a first-encounter enrollment sample is compared with a second-encounter verification sample. Similarly in identification, a new sample is searched against a set of prior enrollments. If the samples are not captured and processed using the same hardware and software, identically configured, the issue of whether the various components are interoperable arises. While interoperability is a desirable and necessary aspect of applications in which multiple vendors sell equipment for capture, processing, and matching, it rests on the availability of well-crafted standards, and specifically, conformity to the various components to those standards. Therefore, sensors might have to conform to imaging specifications, their outputs to image exchange standards, and their transmission might require equipment implementing standardized interfaces. The hazard in biometric applications is that a weak specification or a lack of conformity to a specification might undermine the accuracy of the whole recognition system.

Figure 1 depicts a general interoperable applications. It shows N different biometric capture devices (BCDs) being used to acquire sample data that are then converted from its raw captured biometric data block (cBDB) format into a standardized biometric data block (sBDB) format for enrollment. This is done by any of I template generators. Later, these will be compared by any of K comparison subsystems against verification (or identification) records (sBDBs, in this case) produced from any of J generators processing the output of M BCDs.

This formalism is notional; it defines a five-dimensional component space the last of which is the comparison engine whose outputs support measurements of accuracy. Thus, any combination of five different products can be tested. An interoperable



Interoperable Performance. **Figure 1** Testing the performance of interoperable components.

component is then one that can be used in combination with others. Note that this defines biometric interoperability differently than in some other domains where strict conformance guarantees performance. For example, while a nonconforming implementation of the PGP message standard [1] is likely to give a deterministic and catastrophic failure, a set of fingerprint minutiae automatically extracted from a digital image of an analog trait (i.e., the finger) may well give lower accuracy than those marked by a fingerprint examiner.

This article gives an overview of the biometric interoperability problem and introduces the notion that interoperability should properly be quantified in terms of some relevant performance metrics. It proceeds with examples of interoperability challenges, which motivates the subsequent contribution on interoperability testing.

Interoperability Challenges

Successful recognition depends on the interoperability of all pieces of equipment used in generating both the enrolled and recognition data records. This begins with the acquisition process, and the primary requirement is that the sensors are interoperable. This typically means that the conversion of the analog human trait into the digital sample produces a defined or commonly understood representation of the original. The following paragraphs give examples that undermine interoperability of the three most common modalities.

Fingerprints: It is common in large-scale identity management applications to acquire flat impressions of a subject's fingers, to associate those with a credential, and to verify against one or more of those fingers. While the fingerprint data can be stored in conformity to, for example, the ISO/IEC 19794-4 finger image or 19794-2 finger minutiae standards, subsequent verification attempts depend on the interoperability of the capture devices with the original optical scanner. This is one of the few areas of biometrics where sensors are standardized: The U.S. Federal Bureau of Investigation established a physical imaging specification for optical fingerprint sensors. Known as Appendix F [2], this document regulates the imaging capabilities of the acquisition devices such that the representation of the fingerprint ridge structure is accurately represented in the output image, and that the image is defensible in criminal law enforcement. The specification imposes limits on parameters such as the optical resolution of the device, the amount of geometric distortion, the imaging area, and spatial uniformity.

Face: A face image collected at a distance of 30 cm is unlikely to be interoperable with another acquired at 1 m because of the presence of geometric "fish-eye" distortion. This would affect face recognition systems whose internal representation of the face depends on the relative spatial locations of the various anatomical features. Although a nonlinear re-sampling of the image could correct such distortion, the resulting spatially varying resolution might undermine accuracy. Another possible solution would be to formulate a mathematical representation that is invariant to this kind of distortion. The actual approach from the commercial and user communities has been to regulate the acquisition process via a formal technical standard. This standard ISO/IEC 19794-5:2005 *Face Image Data* requires distortion to be absent, and the amendment ISO/IEC 19794-5/Amd. 1 *Conditions for Taking Photographs for Face Image Data* requires the subject to be positioned at least 0.7m from the camera.

Iris: The interoperability of three iris cameras was measured in the 2005 ITIRT trial [3]. The results of cross-matching images using a single iris recognition package showed that cross-camera accuracy was generally worse than that for single-camera matching. Possible causes for this, which was not asserted in the report, might be the differences in the spectra of the infrared illuminants and in the compression applied post capture.

Interoperable Data Formats

▶ Biometric data interchange standards have been developed to advance interoperability of most of the main biometric modalities. Standards exist for both images and signals, and for "raw" sample data and for processed data. The major extant internationally standardized records are tabulated in Table 1. The standards define a syntactic representation of the data in question. These are usually compact binary encodings of the data suitable for storage on a smartcard or for transmission across a bandwidth-limited communications channel.

Any interoperability problems that could arise from different implementations of the standards might not be revealed until a test is run or a deployment occurs. Although the possible problems are very specific to the standards, the general case is that problems can be expected when two very different sensors are used. For example, in DNA typing, problems would occur if the two sets of loci were disjoint. To avoid such effects, the standards variously regulate the biometric acquisition process.

Interoperability Testing

As various interoperability tests were staged around the world [5, 3, 4], the Working Group 5 of ISO/IEC JTC 1's Subcommittee 37 on Biometrics, which standardizes biometric performance tests, started work on interoperability testing. This culminated in 2007, in ISO/IEC 19795-4 *Interoperability Performance Testing* [6], which establishes procedures for the conduct of tests such as those listed.

The standard requires a testing lab to establish and identify one or more application specific figures of merit, such as false non-match and false match rates, and to report them in the manner presented in Table 2. This shows the interoperability of INCITS 378 fingerprint minutia template [7] generators and matchers. It assumes an enrollment template generated by equipment identified by the row label is later verified against a template generated and matched by equipment from the supplier identified in the column headings. (Please note: The notable results here are that the lowest false non-match error rates are lowest when a single company executes all three functions. That this "native mode" gives better performance than the interoperable cases off the diagonal has been attributed to idiosyncratic minutia placement and selection strategies present in minutia detection algorithms.)

The standard also establishes procedures for how a certification body could use mutually low error rates in an interoperability test as a criterion to identify a core group of interoperable products. The standard then addresses how to maintain a certification program in which products are tested in regular ongoing testing

Interoperable Performance. **Table 1** Biometric data interchange standards for various modalities

Standard	Modality	Processing
19794-2:2005	Fingerprint minutiae	Template
19794-4:2005	Fingerprint	Raw Image
19794-5:2005	Face	Raw or Normalized 2D Image
19794-6:2005	Iris	Raw or Polar Image
19794-7:2007	Signature time series	Multivariate signal
19794-8:2006	Fingerprint skeleton	Processed image
19794-9:2006	Vascular	Raw Image
19794-10:2007	Hand geometry	Binary silhouette Image
19794-13:2010 (est)	Voice	Raw Signal
19794-14:2010 (est)	DNA	Type Signal

campaigns spanning perhaps several years, and in which there might be systematic changes in the difficulty of the test (due to environment, for example).

Sufficiency of a Biometric Data Interchange Standard

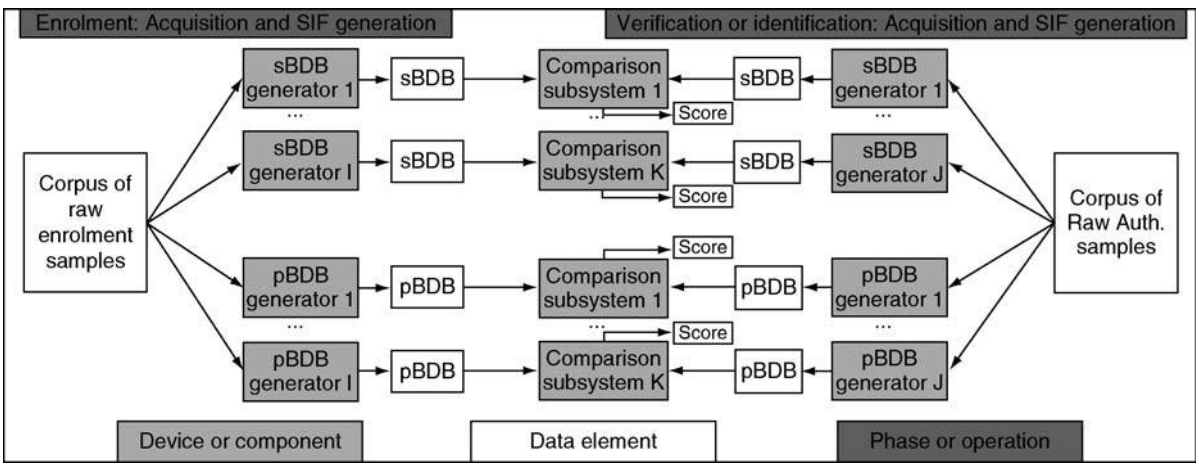
Biometric data interchange standards support interoperability by allowing developers to implement products producing and processing records conforming to a known format. Such standards define the syntactic and semantic representations of the data. For example, a depth value in a 3D face image might be encoded as unsigned integer, but the precision might be commercially important; if one supplier can accurately determine depth to within 0.1 mm, then they would lose accuracy if the standard provides only for depth resolutions of 0.5 mm.

Together, the consensus specifications established in a data interchange format standard might offer less accuracy than a totally unconstrained representation of the biometric data, and the quantification of such a loss goes to the *sufficiency* of a biometric data interchange standard. A point to be noted here is this terminology is that adopted in the international biometric interoperability performance testing standard [6] by answering the question, "does standardized data offer accuracies approaching that of unconstrained, nonstandard representations?".

Figure 2 depicts an offline testing methodology for sufficiency. The MINEX I study [4] quantified sufficiency for the (x, y, θ, type) encoding of the minutiae defined in INCITS 378 [7]. The excerpted results

Interoperable Performance. Table 2 Cross-vendor interoperability for a subset of the minutia detection and matching algorithms evaluated in NIST's MINEX [4] minutia interoperability baseline. The values are false non-match rates at a fixed false match rate of 0.01, for single finger verification on a large offline database

		Verification template and matcher provider					
		A	B	C	D	E	G
Provider of enrollment template	A	0.0136	0.0549	0.0458	0.0225	0.0641	0.0417
	B	0.0218	0.0251	0.0385	0.0173	0.0402	0.0192
	C	0.0357	0.0428	0.0225	0.0204	0.0519	0.0348
	D	0.0207	0.0357	0.0301	0.0140	0.0485	0.0316
	E	0.0236	0.0365	0.0340	0.0225	0.0301	0.0286
	G	0.0300	0.0291	0.0447	0.0205	0.0390	0.0129



Interoperable Performance. Figure 2 Testing the sufficiceny of a data interchnage standard. The samples from a fixed corpus are converted to both proprietary and standardized biometric data blocks (pBDBs and sBDBs, repsectively) and then these are recognized by comparison subsystems from the same suppliers.

Interoperable Performance. Table 3 False non-match rates at fixed rate of 0.01 for fingerprint verification algorithms using fully proprietary and formally standardized templates. The proprietary accuracies would only be available in an interoperable application if the images were exchanged

Kind of template	Provider of template generator + matching algorithm				
	A	B	D	E	G
Proprietary	0.0089	0.0189	0.0089	0.0251	0.0047
Standard	0.0136	0.0251	0.0140	0.0301	0.0129

of Table 3 show that proprietary implementations outperform the standard representation by less than the variation between interoperable pairs observed in Table 2.

Summary

Although interoperability can be supported by placing appropriate specifications on the various components, particularly sensors, biometric recognition performance tests are sometimes the only means of quantifying interoperability. More importantly, accuracy testing is the most operationally relevant measure of interoperability.

Related Entries

► Interoperability
► Performance Testing
► Standards

References

1. Callas J., et al.: RFC 4880 - OpenPGP Message Format (2007). Http://tools.ietf.org/html/rfc4880
2. Hopper T., et al.: IAFIS Image Quality Specifications, EBTS Appendix F. Tech. rep., FFBI Criminal Justice Information Services Division (2008). Http://www.fbibiospecs.org
3. Thieme M., et al.: Independent testing of iris recognition technology final report. Tech. rep., International Biometric Group (2005). Http://www.biometricgroup.com/ITIRT/
4. Grother P.J., et al.: Minutiae Exchange Interoperability Test MINEX - Performance and Interoperability of the INCITS 378 Fingerprint Template. Tech. Rep. NIST 7296, National Institute of Standards and Technology, Gaithersburg, Maryland (2006). Available at http://fingerprint.nist.gov/minex
5. Campbell J., et al.: Seafarers' Identity Documents - Biometric Testing Campaign Report. Tech. rep., International Labour Organization (2005)
6. JTC 1, SC37 Biometrics, Working Group 5: ISO/IEC 19795-4 Interoperability performance testing (2007). http://webstore.ansi.org
7. INCITS M1, Biometrics: INCITS 378:2004 Fingerprint minutia format, 1 edn. (2004). URL http://webstore.ansi.org. American National Standard for Information Technology

Intraclass

Intraclass refers to different instances of the same subject seen under different viewing conditions, illuminations, etc. Ideally, the extracted features should be similar for different instances of the same subject.

► On-Card Matching

Intricated

Intricated is having many complexly arranged elements. A secret key and fingerprint minutia are intricated into a template so that it is not possible to get the key or the minutiae from the template.

► Fingerprints Hashing

Intricated Biometrics

► Fingerprints Hashing

Intrinsic Dimensionality of a Manifold

Intrinsic dimensionality of a manifold is the minimum number of parameters necessary to parameterize a manifold.

► Manifold Learning

Intrinsic Direction of Fingerprint

A unique direction that is defined by the global direction variations of a fingerprint itself. It should be possible to compute it from different fingerprints of the same finger.

► Fingerprint Features

Intrinsic Distance

Intrinsic distance is the length of geodesic between two points on a manifold. It is also referred to as geodesic distance.

► Manifold Learning

Intrinsic Failure

Intrinsic failure is the security lapse due to an incorrect decision made by the biometric system. A biometric verification system can make two types of errors in decision making, namely, false accept and false reject. A genuine (legitimate) user may be falsely rejected by the biometric system due to the large differences in the users' stored template and input biometric feature sets. These intra-user variations may be due to incorrect interaction by the user with the biometric system (e.g., changes in pose and expression in a face image) or due to the noise introduced at the sensor (e.g., residual prints left on a fingerprint sensor). False accepts are usually caused by lack of individuality or uniqueness in the biometric trait, which can lead to large similarity between feature sets of different users (e.g., similarity in the face images of twins or siblings). Both intra-user variations and inter-user similarity may also be caused by the use of non-salient features and non-robust matchers.

► Security Issues, System Design

Invariant–Covariant

A function is invariant under a certain family of transformations if its value does not change when a transformation from this family is applied to its argument. A function is covariant when it commutes with the transformation, i.e., applying the transformation to the argument of the function has the same effect as applying the transformation to the output of the function. A few examples may help to explain the difference. The area of a 2D surface is invariant under 2D rotations, since rotating a 2D surface does not make it any smaller or bigger. But the orientation of the major axis of inertia of the surface is covariant under the same family of transformations, since rotating a 2D surface will affect the orientation of its major axis in exactly the same way. Based on these definitions, it is clear that the so-called local scale and/or affine invariant features are in fact only covariant. The descriptors derived from them, on the other hand, are usually invariant due to a normalization step.

► Local Image Features

Iris

The iris consists of muscle tissue that comprises of a sphincter muscle that causes the pupil to contract and a group of dilator muscles that cause the pupil to dilate. The back surface is covered by a layer of pigmented epithelial tissue. The outer edge is attached to the sclera.

► Anatomy of Eyes
► Iris Image Data Interchange Formats, Standardization
► Template Protection

Iris Acquisition Device

Ryan Rakvic[1], Randy Broussard[1], Lauren Kennell[1], Robert Ives[1], Robert Bell[2]
[1]The John Hopkins University, Applied Physics Laboratory, Annapolis, MD, USA
[2]Biometrics Alliance, Severna Park, MD, USA

Synonyms

Iris camera; Iris image capture device; Iris reader; Iris scanner

Definition

An iris acquisition device is a device that acquires an iris image and compares and matches it to a collection of other iris images. Consumer enthusiasm and technological advancement have fostered the creation of many types of iris acquisition devices that vary in many dimensions. The following essay, surveys and reports iris acquisition devices that exist in the market today.

Introduction

Today, the popularity of iris acquisition devices is gaining momentum mainly due to the fact that they are considered perhaps the most accurate way to identify a human. In addition, iris acquisition devices have the ability to accommodate a large group size, return responses faster than any other security metric, and adapt and utilize a range of advancing technological resources.

Many factors help determine the performance and popularity of an iris acquisition device. The heart of a device is represented by the ▶ iris recognition algorithm, and the resulting accuracy is evaluated with a combination of performance metrics, including false acceptance rate (FAR) and false rejection rate (FRR). The most popular algorithm in use today was invented by iris recognition pioneer John Daugman [1]. The response time of an algorithm is also an important factor in determining the overall popularity of such systems.

Like any young technological product, iris recognition systems have been drastically evolving, driven both by consumer demands and technological improvements. Today, a differentiating factor amongst iris recognition systems is the variance in the product form.

Current iris recognition systems vary in terms of size, weight, ▶ focal distance, and hardware and software portability. The advancement of technology has dramatically impacted iris recognition systems, providing the opportunity to create smaller, more portable autonomous systems. However, the need for larger and more powerful traditional systems still exists. For example, in an airport setting where space and power are not limitations, the primary objective is having the most secure system possible.

The current deployment of iris recognition systems ranges from government to industrial to even private use. Other uses are automated international border crossing, airport and aviation security, database and computer access, hospital access, and countless other private industrial settings. For example, in the UK, five airports are currently using iris recognition systems for security. In the United Arab Emirates, the largest system is employed to prevent deportees from re-entering the country illegally, with over ten billion iris comparisons performed daily. An in-depth list of deployments is given by Daugman [2].

This chapter summarizes the key aspects of the iris recognition systems that are deployed or available for deployment at the present time. The objective here is not to perform a rigorous performance analysis of current iris recognition systems, but to survey and report on information that is publicly available. When specific information is not available, comments are made on the highlights of each iris recognition system. In particular, sufficient information was not always available to report system performance. Additionally, knowledge of the recognition algorithm that is employed by the iris system is not always available. Besides commenting on the key aspects, the current deployment of each iris system, including a representative quotation when appropriate is also highlighted.

Iris Acquisition Devices

The following list of iris acquisition devices is in alphabetical order. Table 1 lists the iris products while giving brief comments about each one. To our knowledge, this information is comprehensive; however there may be some products that are inadvertently not listed below.

Iris Acquisition Device. **Table 1** List of iris acquisition devices

Product	About
Authenticam™ DT120	Highly accurate, portable, affordable, designed for PC authentication purposes, utilizes the Daugman recognition algorithm
BM-ET200	Unit works in standalone mode or can be networked with other security devices, recognition results in 0.3 s, utilizes the Daugman recognition algorithm
BM-ET330	Portable, 5 lbs, 150–183 cm focal distance, dual eye, recognition in less than 1 s provides voice guidance, utilizes the Daugman recognition algorithm
HIIDE™ Series 4	World's first hand-held, small and lightweight 2 lbs, 3 ounces, 8″-10″ focal distance, multi-modal enrollment and recognition device, stand-alone or with a PC, utilizes the Daugman recognition algorithm
Iris642	DSP processor based embedded system, small and lightweight for versatile use
Iris on the Move™	Walk-through portal, extremely low false match rate, ability to identify up to 20 moving subjects per minute, fewer constraints on users than any other iris recognition system
IrisAccess 4000	Intuitive visual user interface; both eyes are captured virtually simultaneously; multi-factor authentication, utilizes the Daugman recognition algorithm
iCAM 4000/4010	Compact, 5 lbs, 10.2–14.2 in. focal distance, low profile, designed with architectural aesthetics in mind, utilizes the Daugman recognition algorithm
iCAM 4100/4110	Compact, 5 lbs, 10.2–14.2 in. focal distance, includes a keypad accepting up to 10 digit PINs affording an additional level of two factor authentication, utilizes the Daugman recognition algorithm
IrisGuard IG-H100	Handheld iris recognition camera with USB interface, 750 g, 12–30 cm focal distance, versatile design, utilizes the Daugman recognition algorithm
IRISPASS-M	Designed to be connected to a PC, 11 lbs, fully automatic, 2-eye, 10–24 in. focal distance, intuitive user interface, voice guidance, identification is complete in less than 1 s after image acquisition, utilizes the Daugman recognition algorithm
I SCAN 2	Durable and compact, held directly to eyes, 1 lb, compatible with known iris matching algorithms
HBOX	Embedded in "through-put environments," IR based, unobtrusive, real time, recognizes at a distance of 1–2 m and in motion at up to 3 m/s
JPC1000	Snaps to your computer, focal distance 15 cm, 80 g
JPC1500	Desktop device, focal distance 30 cm, 400 g
Mobile-Eyes	Hand-held tethered dual-iris capture device, weighs 2.8 lbs, held directly to eyes
Neoris 2000	Standard camera system which incorporates the company's unique stereo camera technology that captures iris and face images at the same time
PIER-T™	Rugged handheld device that allows the operator to both enroll and identify individuals tethered to a host PC or laptop, 12 ounces, utilizes the Daugman recognition algorithm
PIER™ 2.3	Rugged hand-held device, 16.5 ounces, 4–6 in. focal distance, enrollment and identification performed on handheld, utilizes the Daugman recognition algorithm

Authenticam™

The Panasonic "Authenticam™" DT120 is an iris recognition system that is lightweight, compact, and has an iris capture distance of 19–21 in. An example is illustrated in Fig. 1. It utilizes a version of the Daugman [1] recognition algorithm to make this system highly accurate. It is designed to work with a host PC, preventing unauthorized access using what is called "► Private ID™" recognition software.

"The Panasonic Authenticam addresses the core issues of helping an organization diminish or eliminate the costs of password management, reduce the risks of privilege and policy management, and increase the security of a private key. In addition to providing security for information access applications, the iris recognition system camera can be used for video conferencing and online collaboration." [3]

There are multiple product offerings with the base BM [4] Authenticam™ name. The BM-ET200

Iris Acquisition Device. **Figure 1** An example of the Panasonic Authenticam DT120 iris recognition system (from http://www.eyenetwatch.com/iris/panasonic-authenticam.htm).

Iris Acquisition Device. **Figure 2** An example of the Panasonic BM-ET 330 (from http://panasonic.co.jp/pss/bmet330/en/).

[4] unit works in standalone mode or can be networked with other security devices. The claim is that the BM-ET200 has recognition results in 0.3 s. A different model, the BM-ET330, shown in Fig. 2, also provides recognition in less than 1 s while providing voice guidance. The BM-ET330 captures both eyes simultaneously. It can accommodate a range from 150 to 183 cm and weighs approximately 5 lbs. Both models utilize the Daugman [1] recognition algorithm. The performance is scientifically evaluated in an independent study performed by the International Biometric Group [5].

HIIDE™ Series

The HIIDE™ [6] is known as the world's first multimodal hand-held enrollment and recognition device. Illustrated in Fig. 3 is the small and lightweight HIIDE™ Series 4. The Series 4 model weighs only 2 lbs, 3 ounces and the focal distance is 8–10 in. It employs the highly accurate Daugman 2Pi algorithm [1]. Because this series was first designed for the US Department of Defense, it can enroll up to 10,000 biometric portfolios. The HIIDE™ can operate while connected to a PC. It can also operate fully in stand-alone mode, therefore not requiring processing power from a PC. Additionally, it has the ability to connect to USB devices, including live-scan devices, passport or card readers or an external keyboard and mouse.

"The device can operate in extreme and rugged mobile environments, as well as on a desktop connected to a host personal computer or network. This makes it ideal for mobile identification of individuals on the battlefield, at border checkpoints, in airports, in detention centers, and for checking individuals against known watch lists in addition to naval and coast guard applications." [6]

Iris

The Iris642 [7] is the only known DSP processor based embedded system. Because of its embedded nature, it is small, lightweight, and versatile. This system is compatible with third party cameras and can also work in a networked LAN. The Iris642, although a small device, has the ability to do face and iris recognition together. The Iris642 employs a sophisticated algorithm that has resulted in very good recognition scores.

"IriTech uses a variable multi-sector analytic method that selectively utilizes only the good portions of the captured image. Even if the image of the eye is adversely

Iris Acquisition Device. **Figure 3** An example of the L-1 HIDE Series 4 (from http://www.llid.com/images/stories/solutions/hiide_product_sheet.pdf).

affected by eye glasses, contact lenses, tears, eyelids, or eyelashes, IriTech's technology can operate with no discernible performance degradation as long as at least 50% of the image sectors are good at the time of registration and at least 25% are good at the time of identification. Our tests show that the FAR and FRR do not change in any significant way." [7]

Iris on the Move™

As the name indicates, Iris on the Move™ (IOM) [8] provides iris detection and recognition at the "speed of life." Physically, the system permits a person to walk through a portal (shown in Fig. 4), with normal walking speed, for iris detection and capture. It does not require people to stop or to remove their glasses. In addition to being the system with the least constraints, IOM claims to be an extremely good performer with low false match rates and has the ability to capture 20 moving subjects per minute. It also captures and performs recognition on both eyes simultaneously. IOM has many foreseeable applications.

"IOM provides a practical and valuable solution to a critical need for security. Transportation facilities can depend on IOM to expedite safe and secure travel, while secure facilities such as government buildings, courthouses, or power plants can count on its reliable access control to safeguard occupants. From entertainment venues to office buildings, IOM's fusion of security and convenience ensures that the right people are in the right place." [8]

IrisAccess

The IrisAccess product line has been in existence since the late 1990s. All IrisAccess [9] products employ the

Iris Acquisition Device. **Figure 4** An example of the Iris on the Move™ (from http://www.sarnoff.com/products/iris-on-the-move).

Daugman [1] recognition algorithm. Specifically, the IrisAccess 2200 series has been designed for high security server rooms, safety deposit boxes, and other top security areas. As for the level of security, it was purported to be the best at the time of its release. However, the IrisAccess 2200 is no longer serviced, and its successor is the IrisAccess 3000. The performance of the IrisAccess 3000 is scientifically evaluated in an independent study conducted by the International Biometric Group [5]. It is currently deployed by the Albany International Airport.

"After a comprehensive review of existing biometric solutions, the LG 3000 software and hardware platform were chosen for its proven positive identification and authentication of individuals gaining access to secure areas. The technology provides ease-of-use and an accurate audit trail of 'who' opens the doors, not

Iris Acquisition Device. **Figure 5** An example of the IrisAccess 4000 (from http://www.lgiris.com/ps/products/irisaccess4000.htm).

just what card was used to open the doors. The LG Iris solution has been easily integrated into our access control system which handles over 1,000 employees and its performance has promoted [*sic*] us to explore increasing the use of the LG iris identification solution." [9]

The latest and greatest system in this product line is the IrisAccess 4000, illustrated in Fig. 5. The key accessories of the recent release of the 4000 include application versatility, integration flexibility, and enrollment speed. It also captures and recognizes both eyes simultaneously. As with any product chain development, the feature set has become more user-friendly.

"Intuitive visual user interface enables users to quickly position themselves for enrollment or recognition as images of both eyes are captured virtually simultaneously. Audio prompts improve speed of enrollment and recognition performance while a motor-driven auto-tilt mechanism makes adjusting the camera for proper height a simple 'one touch of a finger' proposition." [9]

Included in the IrisAccess 4000 product line is the iCAM series. The iCAM4000/4010 is designed to be compact, weighing less than 5 lbs, and the 4010 model can embed a smartcard. The iCAM4000 is ideal for wall-mount, as it includes a motorized height adjust and face-badging camera. The operating distance is between 10.2 and 14.2 in. The iCAM4100/4110, in addition to the previous features mentioned, also come available with a keypad and LCD display for real-time communication.

"Multifactor authentication can also be delivered by the 16-element keypad that comes standard on the iCAM4100 unit. The authentication options afforded by being able to configure iris authentication by left, right, either or both eyes plus a smartcard token, and in the case of the iCAM4100, a keypad, are simply unmatched by any other iris recognition offering on the market." [9]

IrisGuard

The IrisGuard IG-H100 [10] (see Fig. 6) is a versatile handheld iris recognition camera that provides a USB interface, weighing only 750 g. The focal distance is between 12 and 30 cm. It utilizes the highly accurate Daugman recognition algorithm [1]. It is amenable to accessories that make it ideal for kiosk or wall mount applications. The IrisGuard IG-H100 is utilized around the world in many high-level security venues including the United Arab Emirates (UAE).

"The UAE interior ministry has claimed continuing success for its nation-wide iris-recognition network, which includes the country's airports in its coverage. The system, which uses a single-eye H-100 camera supplied by a UK-based company IrisGuard, was

introduced in 2003 and has since expanded to include 140 iris-recognition stations in 22 enrollment centers and 35 land, sea and air border points across the UAE. The nationwide system has 1.1 million irises stored in its database, according to the interior ministry, and has performed 21 million iris searches for 10.5 million persons over the past 3 years, which makes it the largest search iris database in the world. Using this information, the UAE identified 124,435 individuals who were trying to return to the country illegally with forged documents after deportation." [10]

Iris Acquisition Device. **Figure 6** An example of the IRISGUARD IGH100 (from http://www.irisguard.com/pages.php?menu_id=29&local_type=0).

IRISPASS

The IRISPASS-M [11] (see Fig. 7), in similar fashion to its competitors, provides responses in less than 1 s. It employs the highly accurate Daugman recognition algorithm [1]. The performance of the IRISPASS-WG is scientifically evaluated in an independent study performed by the International Biometric Group [5]. The operational range is approximately 1–2 feet while examining at a height of 57—78 in., and the system weighs 11 lbs. The interface is intuitive as it provides voice guidance. The default model is designed to be connected to a host PC, and accompanying software is available. In the product line are IRISPASS models that connect to mobile phones and personal digital assistants. Like other devices, the IRISPASS-M is employed in law enforcement, border control arenas, and banks.

"Oki Electric Industry Co., Ltd. announced its delivery of IRISPASS®-M iris recognition cameras to Pictet & Cie Bank, a private bank in Geneva, Switzerland. Pictet chose IRISPASS-M to bring the highest level of security at the entrances of high protection room in its newly constructed headquarter office buildings." [11]

ISCAN

ISCAN 2 [12] is a durable (shown in Fig. 8) and compact dual iris capture scanner, weighing just over 1 pound. It is designed for both military and civilian

Iris Acquisition Device. **Figure 7** An example of the IRISPASS-M (from http://www.oki.com/jp/FSC/iris/en/m_features.html).

security programs. Software support is available to provide active mobile enrollment. It is also compatible with many known iris matching algorithms. The scanner is ANSI INCITS 379–2004 and ISO/IEC 19794–6 compliant. The ISCAN technology is also embedded into multi-modal offerings.

"Cross Match will display several other offerings from its comprehensive product portfolio at the BCC, such as its ▶ Multimodal Jump Kits incorporating the ISCAN 2. Jump Kits are used by the military, law enforcement, and other first responders for rapid enrollment as well as for local or remote identification." [12]

HBOX

The HBOX [13] is a revolutionary device that supports a continuous flow of 30 people a minute at up to 3 m per second through a portal like structure. It is multi-modal by recognizing both face and iris. It allows a person to walk through an area while capturing their iris at a distance of 1–2 m. It is a compact device that is 48 × 12 × 12 in., and can be mounted anywhere. An illustration is provided in Fig. 9.

JPC

The Jiris JPC1000 [14] advertises a capability to recognize an eye signature in 1 s and includes software to encrypt or decrypt data. The focal distance is approximately 15 cm and it weights approximately 80 g. The device snaps onto the top of a PC, as illustrated in Fig. 10. The desktop model, the Jiris JPC1500 [14] weights approximately 400 g and has a focal distance of 30 cm.

Mobile-Eyes

Mobile-eyes [15] [16] is a hand-held tethered dual-iris capture device illustrated in Fig. 11. The device weighs 2.8 lbs and incorporates proprietary software.

Iris Acquisition Device. **Figure 9** An example of the HBOX (from www.hoyosgroup.com).

Iris Acquisition Device. **Figure 8** An example of the ISCAN2 (from http://www.crossmatch.com/I_SCAN_2.html).

Iris Acquisition Device. **Figure 10** An example of the JIRIS1000 (from http://www.engadget.com/2006/03/06/jiris-jpc1000-brings-iris-scanning-home/).

Iris Acquisition Device. **Figure 11** An example of the MobileEyes (from www.retica.com).

"They employed new developments in opto-electronics and digital signal processing to create an unprecedented eye biometric system that successfully exploits the full potential of the eye into a four-phase solution that captures, collects, stores, and identifies biometric information. As a transitional or augmented security option, the Retica system integrates iris information into a multi-modal solution that co-manages sequential, tightly-coupled retinal and iris data. Retica's unique patented technology fuses together the images of the retina and the iris, thus significantly facilitating data collection and resulting in higher accuracy." [15]

Neoris

The Neoris 2000 [17] uniquely captures both iris and face images simultaneously with its standard camera system. The uniqueness is provided by its face positioning technology and large focal depth. It is easy to use, and the resulting images are of high quality. The module is based on IriTech proprietary high-performance real-time image capture algorithm.

"Virginia-IriTech, Inc. announced that the company has successfully completed the National Institute of Science and Technology (NIST) independent Iris Challenge Evaluation (ICE) 2006 with excellent results. The NIST ICE tests grant strong independent third party validation to IriTech core iris identification technology. Recently the Neoris 2000 was certified by the Chinese government for sale in China. The Chinese government test required a FAR of less than or equal to 0.0001%, an FRR less than or equal to 0.05% and an enrollment error rate of 0.000%. The NEORIS 2000, running on IriTech algorithms, easily passed these tests, achieving an FAR, FRR, and enrollment error rate of 0.000%. IriTech is the first iris identification company to achieve this Chinese certification." [17]

Iris Acquisition Device. **Figure 12** An example of the PIER 2.3 (from http://www.securimetrics.com/solutions/pier.html).

PIER™

The PIER™-T [18] is a "rugged" hand-held device providing both enrollment and identification functions. Operating tethered to a PC, the device is capable of storing dual iris information for up to 200,000 individuals. The system weighs only 12 ounces, while incorporating state-of-the-art lenses, dual-band illumination, a high-resolution video sensor, and a liquid crystal display screen. The PIER™ performs an image check, and forwards high-quality images to the host PC for recognition. The recognition software employs the industry leading Daugman 2Pi algorithm [1]. This popular and capable system is employed in a wide range of locations, including the Department of Defense (DOD), Biometric Fusion Center (BFC), the Office of Law Enforcement Technology Commercialization (OLETC), and the Space and Naval Warfare Systems Command (SPAWAR).

The PIER™ 2.3, illustrated in Fig. 12, handles iris recognition on the handheld unit. The system weighs

only 16.5 ounces with a focal distance of 4–6 in. It has also been utilized for security purposes.

"The PIER™ 2.3 has been incorporated into the Biometric Application Toolset (BAT), a multi-biometric security platform developed by the US Army, Battle Command Lab. The BAT is being distributed to the Navy, Marines, Army, and other DOD and non DOD agencies. The PIER™ 2.3 is also a critical component in the Rapid Deployment Force "Jump Kit." [18]

Conclusions

Iris acquisition devices continue to grow in popularity, driven mainly by their extraordinarily high accuracy. While the companies that produce these systems have adapted to the desires of their consumers, the ever-changing marketplace has caused these companies to produce a wide range of products that vary in many dimensions. As can be seen with the list above, current iris recognition systems vary in terms of accuracy, size, weight, focal distance, form factors, hardware and software portability, and most importantly, current deployment. As computer and camera technology continues to improve and consumer demand continues to evolve, it is expected that future iris recognition systems will become more accurate, smaller, and more portable causing a large increase in future deployment.

Related Entries

- ▶ Iris Device
- ▶ Iris on the Move™
- ▶ Iris Recognition System Deployments
- ▶ Overview
- ▶ Performance Evaluation
- ▶ Performance Measures
- ▶ Performance Testing Methodology
- ▶ Standardization

References

1. Daugman, J.: How iris recognition works. In: Proceedings of the IEEE International Conference on Image Processing (ICIP), Rochester, New York (2002)
2. Daugman, J.: 17 December 2007. http://www.cl.cam.ac.uk/~jgd1000/
3. Eyenetwatch.: 17 December 2007. www.eyenetwatch.com/iris/panasonic-authenticam.htm
4. Panasonic: 17 December 2007. www.panasonic.com/business/security/home.asp
5. International Biometric Group: 17 December 2007. www.biometricgroup.com/reports/public/ITIRT.html
6. *L-1*: 17 December 2007. http://www.l1id.com/images/stories/solutions/hiide_product_sheet.pdf
7. Iritech: 17 December 2007. www.iritech.com/product_1-1.htm
8. Sarnoff: 17 December 2007. www.sarnoff.com/products/iris-onthe-move
9. LG Electronics: 17 December 2007. http://www.lgiris.com/ps/products/index.htm.
10. Irisguard: 17 December 2007. www.irisguard.com
11. OKI: 17 December 2007. http://www.oki.com/jp/FSC/iris/en/
12. Crossmatch: 17 December 2007. www.crossmatch.com/I_SCAN_2.html
13. Hoyos Group: 17 December 2007. www.hoyosgroup.com
14. Retica: 17 December 2007. http://www.retica.com/site/company/index.html
15. Jiris: 17 December 2007. www.jiristech.com
16. Morse, B.G.: CEO Retica Systems, "Product Availability." Email to Ryan Rakvic. 26 November 2007
17. Iritech: 17 December 2007. http://www.iritech.com/
18. L-1: 17 December 2007. www.l1id.com

Iris at a Glance

- ▶ Iris on the Move™

Iris Biometric

- ▶ Iris Recognition, Overview

Iris Camera

- ▶ Iris Acquisition Device
- ▶ Iris Device

Iris Capture

► Biometric Sample Acquisition

Iris Data Interchange Standards

► Iris Image Data Interchange Formats, Standardization

Iris Databases

Damon L. Woodard[1], Karl Ricanek[2]
[1]Clemson University, Clemson, SC, USA
[2]University of North Carolina Wilmington, Wilmington, NC, USA

Definition

An iris database is a collection of images that contain, at a minimum, the iris region of the eye. The images are typically collected by sensors that operate in the ► visible spectrum, 380–750 nm, or the near infrared spectrum (NIR), 700–900 nm. The visible spectrum image can be stored as a color image or as an intensity image. The NIR image is always stored as an intensity image.

Introduction

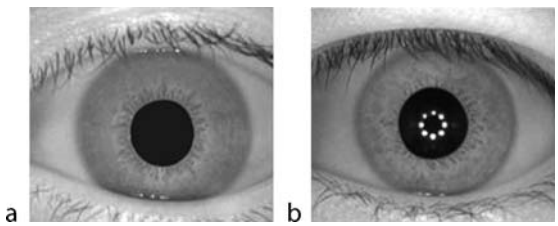
Successful biometric research requires the analysis of human data. For biometric researchers to demonstrate the effectiveness of proposed iris segmentation/recognition techniques and allow fair comparisons with existing methods, publicly available iris databases are required. The perfect iris-image database should be sufficiently large, consist of images collected from a large and heterogeneous group of subjects, and contain images that depict noise factors typically encountered in real world applications. In the following sections, several publicly and freely available iris-image databases are described.

CASIA V1.0, V2.0, IrisV3 Databases

The CASIA Iris Image databases were collected by the Institute of Automation at the Chinese Academy of Sciences. The original CASIA database (CASIA V1.0) is one of the oldest publicly available systems for evaluation of the iris biometric modality; hence, CASIA V1.0 has been widely used for research and evaluation. The database [1] consists of 756 320 × 280 intensity iris images of 108 eyes captured using a sensor developed in-house. The images are stored as 8 bit gray-level JPEG files. Seven images are captured from each eye during two sessions, three during the first session and four during the second. The data acquisition environment was highly constrained which limited the noise types present in the image to iris occlusion from eyelids and eye lashes. It was determined that the iris images of the CASIA V1.0 database had been photographically edited by replacing the pupil region with a circular region of uniform intensity [2], which is illustrated in Fig. 1. Therefore, this version of CASIA should not be used for algorithm development or evaluation.

The CASIA V2.0 database was released later and contains 2,400 images collected from 60 eyes. Twenty images per eye were collected using two different sensors, the in-house sensor and the OKI Iris Pass system. In order to receive copy of CASIA V2.0, CASIA V1.0 must be downloaded first.

The CASIA-IrisV3 database consists of 22,051 images captured from 1,500 eyes of more than 700 subjects, organized three separate databases of disjoint subject groups [3]. The "Interval" subset contains 2,655 images captured from 396 eyes of 249 subjects. The 320 × 280 intensity images were captured indoors using a self-developed sensor that was used in CASIA V1.0 during two acquisition sessions over at least a



Iris Databases. Figure 1 Example images from CASIA V1.0 and CASIA-IrisV3 "Interval" databases. (**a**) An image from CASIA v1.0 (**b**) An image from CASIA-IrisV3.

1 month time. Included in the "Interval" database is an unedited – the pupil region was not replaced to mask the NIR illumination pattern – superset of the CASIA V1.0 database. An example of the unedited image can be found in Fig. 1(b).

The "Lamp" database contains 16,213 images captured from 819 eyes of 411 subjects. The 640 × 480 grayscale images were captured indoors with varying illumination with the OKI IRISPASS-h sensor during a single acquisition session. The "Twins" database contains 3,183 images captured from 400 eyes of 200 subjects. The 640 × 480 intensity images were captured outdoors using the OKI IRISPASS-h sensor during a single acquisition session.

UPOL Database

The UPOL iris-image database was constructed at the University of Palackého and Olomouc in the Czech Republic [4, 5]. The database contains 384 768 × 576 color images captured from 128 eyes of 64 subjects (three images per left and right eye). The images are stored as 24-bit color images in the PNG image format. Unlike the other iris-image databases, the UPOL database was acquired using an optometric sensor (TOPCON TRC50IA optical device connected to a SONY DXC-950P camera). As a result, the images are of very high quality and are practically noiseless, as shown in Fig. 2.

BATH Database

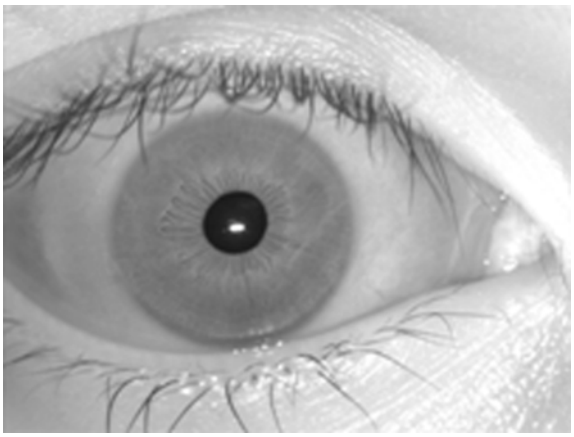
The University of Bath, UK, iris-image database is composed mainly of images captured from its ethnically diverse staff and students [6]. The database contains over 16,000 1280 × 960 iris images collected from each eye of 800 subjects. A commercial version of this database claims to be twice as large. The sensor used for data collection was the ISG LightWise LW-1.3-S-1394. Illumination was provided using an array of infrared LEDs positioned such that reflections were confined to the pupil area within the image, which is illustrated in Fig. 2. An infrared pass filter was used to reduce environmental light reflections. The researchers were able to achieve high quality images using this framework. The main sources of noise within the images are due to the occlusion of the iris by eyelids and eye lashes (Fig. 3).

ICE2005 and ICE2006 Databases

The Iris Challenge Evaluation (ICE) was conducted and managed by the National Institute of Standards and Technology (NIST) and consisted of an iris recognition challenge problem distributed to various participating universities, government agencies, and biometric technology companies [7]. The broad goals of ICE were to facilitate the development of iris recognition technology along with assessing the state of iris recognition systems.

The ICE2005 database consists of 2,953 640 × 480 intensity images with varying numbers of images acquired from each subject [8]. The images are 8-bit gray-level in the TIFF image format. Exactly 1,425 images

Iris Databases. Figure 2 Example images from UPOL iris image database.



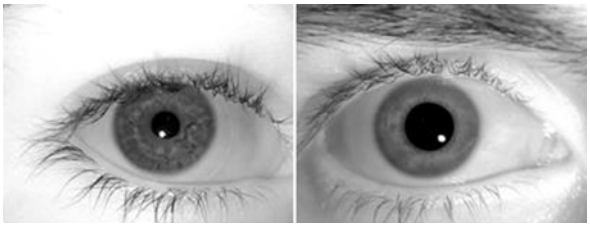
Iris Databases. Figure 3 Image from the Bath iris-image database demonstrating the illumination array centered in the pupil [6].

were captured from the right eye of 124 subjects and 1,528 images were captured from the left eye of 120 subjects. Images of both eyes of 112 subjects were collected resulting in 132 total subjects used to construct the database. Images were collected using an LG Iris Access 2200 iris camera. There are representations of various types of image noise within the database such as iris occlusion, poor focus, and partially captured eyes as illustrated in Fig. 4. A larger database of over 65,000 images captured from 356 subjects in planned for release. The new database is a superset of the ICE2005 database and the later constructed ICE2006 database.

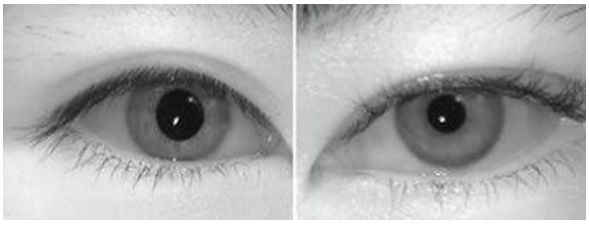
MMU1 and MMU2 Databases

The Multimedia University constructed a relatively small data set of 450 320 × 240 grayscale iris images designated as MMU1 [9]. These images were captured using the LG IrisAccess 2200 camera. Examples images are shown in Fig. 5. Information regarding the number of eyes and subjects used to construct this database is not provided.

A new database of 995 iris images captured from 100 subjects from Asia, Middle East, Africa, and Europe, using the Panasonic BM-ET100US Authenticam. Five images were taken from each of the subjects' eyes. Five left eye images were discarded due to a subject's diagnosis of cataract disease. As with many of the previous databases, noise present in the images was mainly due to iris occlusion.



Iris Databases. **Figure 4** Example images from ICE iris database.

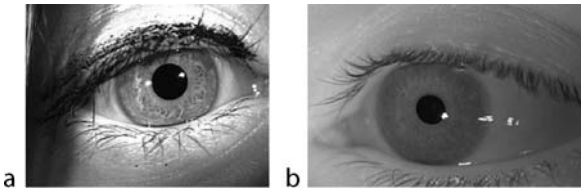


Iris Databases. **Figure 5** Example images from MMU1 iris database.

WVU Non-Ideal Iris and Off Axis/Angle Databases

The West Virginia University constructed two iris-image databases [10, 11]. The iris portion of the database consists of 3,099 480 × 640 grayscale iris images captured from 244 subjects. The images were captured during a single session using the OKI IrisPass-H sensor. The number of images from each eye varies between three and six. This database was constructed as a non-ideal iris-image database, and therefore, contains a considerable number of images that depict various types of noise, which could be encountered in a real world scenario. Noise types include iris occlusions, varying illumination, poor focus, and off-angled images. An example image from this database is shown in Fig. 6(a).

Most iris databases are composed of frontal view iris images. The second WVU iris database is composed of images of irises taken at various angles. The database was captured using two sensors. The first set of images was captured using the Sony Cyber Shot DSC F717 camera used in infrared mode. There are 268 2560 × 1920 RGB images captured from 19 subjects during a single acquisition session. During data collection the camera was positioned at the angles of 0, 15, and 30 degrees. The second set of images for this database was captured using a monochrome camera.



Iris Databases. **Figure 6** Example images from WVU non-ideal iris image and off-axis/angle iris image databases. (**a**) Non-ideal iris image example (**b**) Off-angle iris image example.

There are 597 720 × 480 intensity images captured from 73 subjects during a single acquisition session. As with the other image set, the camera was positioned at the angles of 0, 15, and 30 degrees during data collection. An example iris image captured at 15 degrees off-angle is shown in Fig. 6(b).

UBIRIS.v1 and UBIRIS.v2 Database

In order for researchers to test the robustness of iris segmentation/recognition algorithms when using noisy images, University of Beira Interior constructed an iris database of noisy iris images [12, 13]. The images were captured using the Nikon E5700 camera. There are 1,877 2560 × 1704 RGB images captured from 241 subjects during two acquisition sessions. To introduce noise to the process, the location of the acquisition session was changed to facilitate changes in natural luminosity, contrast, reflections, and focus, as shown in Fig. 7.

Recently, the University of Beira Interior made available a second version of the UBIRIS database, UBIRIS.v2, which currently consists of 11,000 images and is only available to participants in the ▶ Noisy Iris Challenge Evaluation (NICE) Part I contest [14].

Summary

Iris-image databases are crucial to the development and advancement of iris-based biometrics. These databases along with prescribed evaluation methodologies allows for direct comparison of iris segmentation/recognition algorithm performance. The databases will increase in size and complexity of iris-image until all algorithmic problems, inefficiencies, and shortcomings have been fully addressed.

Iris Databases. Figure 7 Example images from UBIRIS.v1 iris database [12].

Related Entries

- ▶ Iris
- ▶ Iris Acquisition Device
- ▶ Iris Challenge Evaluation
- ▶ Iris Image Quality
- ▶ Iris Sample Synthesis

Bibliography

1. Institute of Automation, Chinese Academy of Science: CASIA v1.0 Iris Image Database, 2008. http://www.nlpr.ia.ac.cn/english/irds/irisdatabase.htm. Accessed 27 Dec, 2008
2. Phillips, P.J., Bowyer, K.W., Flynn, P.J.: Comment on the CASIA version 1.0 Iris Dataset, IEEE Trans. Pattern Anal. Mach. Intell. **29**(10), 1869-1870 (2007)
3. Institute of Automation, Chinese Academy of Science: CASIA v3.0 Iris Image Database, 2008. http://www.nlpr.ia.ac.cn/english/irds/irisdatabase.htm. Accessed 27 Dec, 2008
4. Dobeš, M., Machala, L.: UPOL Iris Image Database, 2008. http://phoenix.inf.upol.cz/iris/. Accessed 27 Dec, 2008
5. Dobeš, M., Machala, L., Tichavský, P., Pospíšil J.: Human Eye Iris Recognition Using the Mutual Information. Optik **115**(9), 399–405 (2004)
6. University of Bath: University of Bath Iris Image Database, 2008. http://www.bath.ac.uk/eleceng/research/sipg/irisweb/index.html. Accessed 27 Dec, 2008
7. National Institute of Standards and Technology: Iris Challenge Evaluation (ICE), 2008. http://iris.nist.gov/ICE/. Accessed 27 Dec, 2008
8. Liu, X., Bowyer, K.W., Flynn, P.J.: Iris Recognition and Verification Experiments with Improved Segmentation Method. In Proceedings of Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID). Buffalo, NY, 17–18 October 2005
9. Multimedia University: MMU1 and MMU2 Iris Image Databases, 2008. http://pesona.mmu.edu.my/~ccteo. Accessed 27 Dec, 2008
10. West Virginia University: West Virginia University Biometric Dataset Collections, 2008. http://www.csee.wvu.edu/~simonac/CITeR_DB. Accessed 27 Dec, 2008
11. Ross, A., Crihalmeanu, S., Hornak, L., Schuckers, S.: A Centralized Web-Enabled Multimodal Biometric Database. In Proceedings of the 2004 Biometric Consortium Conference (BCC), Arlington, VA, September 2004
12. Proença, H., Alexandre, L.: UBIRIS: A Noisy Iris Image Database. In: Proceedings of the 13th International Conference on Image Analysis and Processing (ICIA2005), Vol. 1, pp. 970–977, 2005
13. SOCIA Lab – University of Beira Interior: UBIRIS.v1 Iris Image Database, 2008. http://iris.di.ubi.pt/ubiris1.html. Accessed 27 Dec, 2008
14. SOCIA Lab – University of Beira Interior: Noisy Iris Challenge Evaluation – Part I, 2008. http://nice1.di.ubi.pt/

Iris Device

James R. Matey
Electrical and Computer Engineering Department,
US Naval Academy, Annapolis, MD, USA

Synonyms

Iris camera; Iris image capture device; Iris reader; Iris camera; Iris scanner

Definition

An iris device is a device that acquires images of the iris for use in biometric recognition. More commonly referred to as iris cameras, readers or scanners, these devices typically include some form of active ▶ near infrared illumination (NIR), since the current iris recognition algorithms were designed for NIR images of the eye. In addition, the device may include features to aid subjects in properly aligning their eyes in the field of view, spoofing countermeasures, and on board processing of the acquired images.

Introduction

Iris recognition is one of the strongest biometrics available [1–4]. Iris recognition is a strong biometric because: (1) the human iris is a complex structure with a high degree of randomness; (2) the iris is protected; (3) the iris is accessible; and (4) the structures of the iris that are used for iris recognition are stable, from early childhood on – in the absence of illness or injury that disrupts the iris tissue.

The first mention of iris patterns as a biometric was likely a paper by Bertillon [5]; several others subsequently suggested iris patterns as a biometric and the idea was a plot element in the 1983 James Bond film *Never Say Never Again* [6]. However, it was not until the early 1990's that John Daugman developed a practical algorithm for iris recognition based on Gabor wavelets [7]. Minor variants on the Daugman algorithm remain the dominant algorithms in commercial iris recognition systems as of 2008, though there are vigorous research efforts into alternative algorithms [8]. The commonly used name for the Daugman algorithms in current use is ▶ iris2pi.

Further discussion of algorithms may be found in other entries that cover the algorithmic aspects of iris recognition in detail. This entry is concerned with devices that can acquire images suitable for iris recognition. The reader may well ask, "What is a suitable image?" That question depends on the algorithms used for recognition, and the determination of the minimum quality image for a given application is still a matter for research. However, ▶ ANSI and ▶ ISO have published standards [9] for iris image quality that are generally accepted for commercial applications. From the standpoint of iris image capture device design, the most important of these image quality metrics are

- ▶ Resolution: 100–200 ▶ pixels across the iris
- Signal to noise ratio: 40 dB, 100:1. For 8 bit pixels exercised over the full dynamic range, the noise would be $\sim \pm 2$ digital numbers (DN). Equivalently, 7 bits should be meaningful.
- ▶ Contrast: For 8 bit pixels, 90 gray levels of separation between the iris and sclera and a 50 gray levels of separation between iris and pupil
- Operating ▶ wavelength: 700–900 nm

Brief History of Iris Image Acquisition Devices

The first commercial iris image acquisition device was the System 2000 from IrisScan [10] in 1995. In the past decade numerous iris image acquisition devices have been introduced, as can be seen from the partial listing in Table 1. To the author's knowledge, there has not yet been a comprehensive, publicly available evaluation of the relative merits of these devices. One of the most comprehensive tests to date was conducted by the International Biometrics Group (IBG) for the US government. There have also been tests of iris recognition algorithms – rather than the acquisition devices – such as the Iris Challenge Evaluation (ICE) competition sponsored by the NIST.

The IBG study compared one product each from Panasonic, OKI and LG. A subsequent IBG study [11] reported on the IrisGuard H100. The lack of more comprehensive iris device evaluations is understandable. Iris recognition works quite well; its failure rates are low. Hence, statistically significant

Iris Device. **Table 1** Partial List of Iris Image Acquisition Devices

Vendor	Model	Year introduced
IrisScan	System 2000	1995
OKI	IrisPass®-S	1998
LG	2200	1999
Sensar	R1	1999
Iridian	Authenticam™	2000
Panasonic	BM-ET-100 – Authenticam™	2001
LG	3000	2002
OKI	IrisPass®-WG	2002
Panasonic	BM-ET-500	2002
IrisGuard	H-100	2003
Securimetrics	Pier™ 2.2	2003
Securimetrics	Pier™ 2.3	2003
OKI	IrisPass-H	2004
Panasonic	BM-ET-300	2004
LG	4000	2005
OKI	IrisPass®-M	2005
Sarnoff Corporation	Iris On the Move™	2005
Securimetrics	HIIDE™	2005
IriTech	Neoris 2000	2006
Jiris	JCP1000	2006
Panasonic	BM-ET-330	2006
Hoyos	Hbox™	2007
Panasonic	BM-ET-200	2007
IrisGuard	AD-100	2008

probes of the false match, false non-match, failure to enroll and failure to acquire rates require many individual trials, making such tests difficult and expensive. The addition of evaluations of ease of use and suitability for various scenarios only compounds the problem. Another non-technical issue is also important: the iris recognition marketplace is intensely competitive – so much that some of the rivals have had protracted legal battles in the courts. Under these conditions, convincing a vendor to participate in a public test whose results might show that their product is in some way inferior to a competitor's product can be far more difficult than any technical challenge. There have been unpublished evaluations of iris devices by government and private industry – some of which may be obtained by request to the device manufacturers.

This entry will trace the operation of a generic iris image acquisition device – from ▶ photons to identity – and discuss the metrics that are important for evaluation of such devices in generic deployment scenarios.

Photons to Iris Image

Iris recognition starts with photons of light from ambient illumination or from an active illuminator within the iris recognition system. The photons impinge upon the subject iris and are partially reflected. The reflected photons fall into two categories, specular ▶ reflection and diffuse reflection. The specular reflection is due to the impedance mismatch between eye tissue and air; the physics is the same as that for the reflections that you see in a piece of otherwise clear window glass or a bathroom mirror. Since the eye is curved, it acts as a convex mirror and reduces the apparent size of objects seen in its specular reflections (Some automobile rear

view mirrors are convex and carry the warning "Objects may be closer than they seem").

The diffusely reflected fraction is the ▶ albedo of the iris and is of the order of 10% in the near infrared (NIR). The diffusely reflected fraction varies across the iris – that is what gives rise to the patterns seen in an iris image. The diffusely reflected photons are scattered in all directions. The lens of the iris camera will intercept a small fraction of the scattered photons; the fraction is, to first order, the area of the lens divided by the area of a half-sphere centered on the subject with a radius equal to the camera-to-subject ▶ standoff distance. Mathematically, the fraction is

$$\frac{1}{2}\left(\frac{r_{\text{lens}}}{x_{\text{camera}-\text{subject}}}\right)^2,$$

where r is the lens radius and x is the camera to subject distance. A lens 1 cm in diameter at a distance of 10 cm will capture approximately 0.1% of the light diffusely scattered from an iris (or any other object).

The lens will focus an image of the iris, and the variation of its diffuse reflection, onto an imaging sensor. The sensor could be a piece of photographic film. In practice, for iris imaging devices it is almost always a piece of silicon in the form of either a ▶ CCD imager or a ▶ CMOS imager. The most important characteristics of the imager are its pixel size, the number of pixels, and the ▶ quantum efficiency (QE) of the pixels.

The pixel size of the imager sets the ▶ focal length of the lens. The iris is approximately 1 cm across for most of the population. ANSI standards require 100 to 200 pixels across the iris; this corresponds to a pixel size at the iris of the order of 100 microns. The pixel size at the imager is of the order of 10 microns. Hence, the lens system of the camera must magnify (minify) the iris by a factor of approximately 0.1. A desired resolution at the iris and a known imager pixel size define the ▶ magnification required. The magnification and the camera-to-subject standoff distance then specify the focal length of the camera lens through the simple lens equation

$$M = \frac{q}{p}$$

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q},$$

where M is the magnification, f is the focal length, and p and q are the subject-to-lens and lens-to-sensor distances respectively. For a magnification of 0.1 and a standoff distance of 2 meters, the lens will have a focal length of ~200 mm. Note that the smallest possible lens-to-sensor distance, q, is f. Hence these considerations also drive the minimum size of the camera/lens package.

The number of pixels in the imager sets the field of view of the imager. For 100 pixels across the iris, the minimum is 100 pixels across the imager – assuming perfect alignment of the subject iris with the camera. In practice, rather more is required. Cameras designed for 200 pixels across iris frequently use imagers that are 640 × 480 – well more than twice the minimum number of pixels. Cameras designed for large standoffs, such as the Iris on the Move™ system, use much larger (2048 × 2048) imagers to relieve the alignment requirements on both the subject and the system.

The quantum efficiency of the sensor is the fraction of photons that are converted into electrons. More electrons give a larger signal and better ▶ signal to noise ratio (SNR). The number of electrons is just the quantum efficiency times the number of photons delivered to the sensor pixel. The primary noise sources for sensors used in iris systems are ▶ read noise and ▶ shot noise. Read noise is introduced each time a pixel is read; the read noise power is to first order constant for a given camera configuration. Shot noise is the result of electron/photon statistics – the randomness with which photons arrive and are converted to electrons. Shot noise is proportional to the square root of the number of electrons. Iris sensors normally operate in a shot noise dominated regime, so doubling the quantum efficiency of the sensor will improve the SNR by the square root of 2, rather than 2. The number of photons depends on both the rate of delivery and the shutter time of the camera. Matey et al published an analysis of the tradeoffs between SNR and other systems design decisions [12].

If there were no concern about the safety of the subject, the SNR of the system could be made arbitrarily large by simply increasing the illumination level on the subject iris. However, the eye is sensitive to illumination – even illumination that it cannot see. The American Conference of Governmental Industrial Hygienists (ACGIH) provides guidelines, ▶ threshold limit values (TLVs), for acceptable exposures to

infrared illumination [13]. The guidelines cover both the irradiance (W/cm^2) at the eye as well as the radiance ($W/sr\text{-}cm^2$) of the source. Any iris imaging system needs to take these guidelines as well as the various national and international safety standards and regulations into account.

In summary, photons travel from the illuminator to the eye and are reflected, specularly and diffusely, to the camera lens. The camera lens focuses the photons onto the camera sensor; the sensor converts the photons to electrons; the number of electrons is measured and the measurement is converted to a digital signal that is then assembled into a digital image for subsequent processing by a biometric recognition algorithm that can determine if the eye in front of the camera has previously been enrolled in the database against which current eye is being compared.

Iris Image Acquisition Device System Metrics

Resolution, SNR, contrast and operating wavelength and the resulting image quality are crucial metrics for the iris images, as noted in the introduction. However, iris image acquisition devices must fit into systems deployed in the real world – and in the real world, those metrics are not enough.

Ease of use, robustness, reliability, ► interoperability and cost are several of the real world systems metrics that are important considerations for iris acquisition devices. Each of these could be the topic of an entry on its own. The article now briefly discusses ease of use.

In the author's opinion, the primary factors for ease of use are ► capture volume, ► residence time and subject motion. Capture volume is the volume over which a good quality iris image can be reliably captured. Small volumes make it difficult for subjects to present their irises to the system. Residence time is the length of time that a subject must hold their iris within the capture volume. Small volumes with long residence times are particularly difficult. Large volumes with short residence times are almost always better.

Subject motion is complicated. How much motion is tolerable depends on direction, longitudinal or transverse to the camera line of sight. Tolerance to motion can be improved by using short shutter times, at the expense of SNR. In general, systems that allow for subject motion are easier for the subjects to use.

Ease of use can be traded off in some circumstances. If the users of a system are well trained and habituated, the system can be successful, even if it is very difficult for a first time user.

Related Entries

► Biometric Data Interchange Format, Standardization
► Biometric System Design
► Iris Encoding and Recognition
► Iris Image Quality
► Iris Recognition, Overview

References

1. Mansfield, A., Kelly, G., Chandler, D. Kane, J.: "Biometric Product Testing: Final Report" (CESG Contract X92A/4009309), Centre for Mathematics and Scientific Computing, UK National Physical Laboratory (2001)
2. Daugman, J.: "How Iris Recognition Works", IEEE Trans. Circ. Syst Video Tech. **14**(1), (2004)
3. International Biometrics Group, "Independent Testing of Iris Recognition Technology, Final Report, May 2005", NBCHC030114/0002. Study commissioned by the US Department of Homel and Security
4. Jain, A.K.: Iris recognition. In: Jain, A.K., Flynn, P.J., Ross, A. (eds.) Handbook of Biometrics, Springer, New York (2007)
5. Bertillon, A.: "La couleur de L'Iris" Annales de Demographie Internationale **7** 226–246 (1886)
6. Never Say Never Again [motion picture], Schwartzman, J. producer. Warner Bros. 1983
7. Daugman, J.: Biometric Personal Identification System Based on Iris Analysis. U.S. Patent No. 5,291,560 (1 March 1994)
8. Phillips, P., Jonathon, W., Scruggs, T., Alice J.: O'Toole, Flynn, P. J., Bowyer, K. W., Schott, C. L., Sharpe, M.: "FRVT 2006 and ICE 2006 Large-Scale Results." NISTIR 7408, March 2007. See also http://iris.nist.gov/
9. ANSI INCITS 379-2004. Iris Image Interchange Format, American National Standards Institute, Inc., New York, NY
10. IrisScan and Sensar were merged to form Iridian. Iridian was subsequently bought by Viisage (now L1 Identity Solutions) and merged into their Securimetrics division
11. International Biometrics Group, "Comparative Biometric Testing, Round 6 Public Report", September 2006
12. Matey, J.R., Ackerman, D., Bergen, J., Tinker, M.: Iris recognition in less constrained environments. In: Ratha, N. and Govindaraju, V. (eds.) Advances in Biometrics, Sensors, Algorithms & Systems, Springer, London (2008)
13. ACGIH, 2006 TLVs and BEIs, ISBN 1-882417-62-3, www.acgih.org

Iris Digital Watermarking

Nick Bartlow, Nathan Kalka, Bojan Cukic, Arun Ross
West Virginia University, Morgantown, WV 26506, USA

Synonyms

Biometric Watermarking; Digital Watermarking

Definition

Iris digital watermarking is a specific type of ▶ digital watermarking that involves the imperceptible embedding of data (the watermark) in iris images (the host) in order to impart additional security to an iris biometric system. The watermark data can be strings of random digits; system specific identifiers, such as organization names and file creation dates; or biometric feature vectors. The additional security offered by the watermark may result from using it as a mechanism for proving the authenticity of the host image, tracking the chain of custody, or incorporating a multimodal biometric option. Iris digital watermarking or ▶ biometric watermarking, in general, is typically used in tandem with ▶ cryptography, and importantly, provides a layer of security that remains intact after the decryption process. A functional iris digital watermark should not degrade the performance of the host biometric system it protects.

Introduction

Coined by Tirkel et al. [1], the term "digital watermark" originated in 1993. Unlike their physical predecessors (i.e., currency, copyright marks, etc.), digital watermarks are usually imperceptible to the human eye, requiring the use of machines for detection and extraction from the host media in which they are embedded. Digital watermarking is closely related to the field of ▶ steganography where secret messages are clandestinely embedded in larger, unrelated messages [2]. The benefits of digital watermarking are somewhat broad and can fall under one or more areas of interest. Traditionally, digital watermarking has been employed as a form of copyright protection that allows an individual to prove (or disprove) ownership by embedding and extracting data suitable for verification. Additionally, digital watermarking can be applied to verify the authenticity of digital media, provide copyright protection or reproduction management, and offer another mechanism for content description [3, 4]. Biometric watermarking is a particular case of digital watermarking where the content of the watermark or the host data (or both) are biometric entities. This imparts an additional layer of authentication to the underlying system. Figure 1 shows examples of the four main classes of digital watermarking.

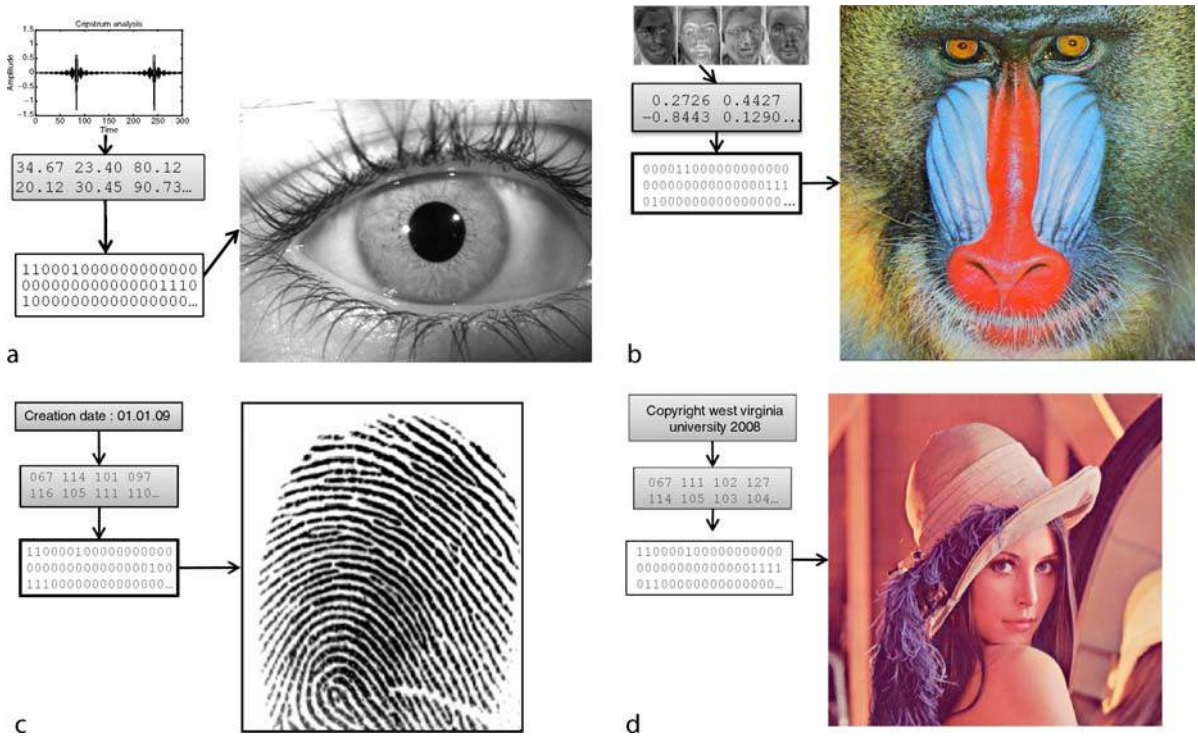
A digital watermark can be characterized using attributes such as visibility, blindess, and symmetry. Kutter provides a comprehensive description of these attributes athttp://www.watermarkingworld.org/. Formalized definitions of these attributes are outlined in Table 1. The issue of visibility or perceptibility relates to whether or not the watermark is noticeable by humans. The issue of blindness indicates whether or not the process of detection and extraction of the watermark relies on the original host data or other auxiliary data. Auxiliary data in semiblind systems can include information dealing with encoding module input parameters, encoding locations, or any other information used to assist in the detection and or extraction processes. Finally, watermarking systems must make use of keys, either private similar to symmetric cryptographic systems or public key pairs akin to asymmetric cryptosystems. Although biometric watermarking system characterized by any combination of these attributes is plausible, an invisible, blind, and asymmetric system is arguably the most difficult to conceive.

The quality of a biometric watermarking system can be evaluated based on five measures: imperceptibility, robustness, fragility, capacity, and performability. Table 2 describes each of these five measures in depth. Although the characteristics imperceptibility, robustness, fragility, and capacity apply to digital watermarking, performability is specific to biometric watermarking.

Algorithms

Overview

A generalized iris digital watermarking framework can be broken into three main operating modules:



Iris Digital Watermarking. Figure 1 The four classes of digital watermarking (**a**) Biometric watermark embedded in biometric host data; (**b**) Biometric watermark embedded in nonbiometric host data; (**c**) Nonbiometric watermark embedded in biometric host data; (**d**) Nonbiometric watermark embedded in nonbiometric host data. Cases (**a**)–(**c**) are examples of biometric watermarking. The images of Lena and the Mandrill are reproduced from the USC-SIPI Image Database.

Iris Digital Watermarking. Table 1 Description of digital watermarking types applicable to biometric watermarking

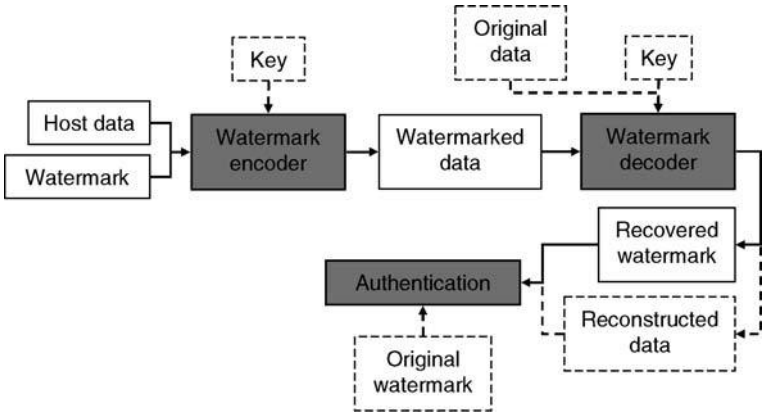
Type	Description
Visible	The embedded information is noticeable by humans, either visually in pictures or audibly in sound files
Invisible	The embedded information is either completely imperceptible or not easily noticeable by humans without the assistance of machines
Public (blind)	The original host file is not required to detect/extract the embedded watermark
Semiblind	The embedded watermark is detected with additional information relative to the watermark encoding scheme, but does not require the entire original host file
Private (nonblind)	The embedded watermark can only be detected/extracted with both the watermarked image and the original host file
Symmetric (Private Key)	A secret / private key is utilized to encode the watermark in the host image. This requires communication of the secret key between the sender and the receiver
Asymmetric (Public Key)	A public–private key pair is used to encode the watermark in the host file. The use of a public key pair prevents the need to communicate a secret key between sender and the receiver, as only the public portions of the key pairs need to be available. Optionally, this method can ensure image integrity and nonrepudiation of origin

encoding, decoding, and authentication (see Figure 2). The first operating module, the watermark encoder, embeds the watermark into the host data. The watermark can range from a random binary bit sequence to biometric eigen-face coefficients utilized for face recognition. Additionally, the encoding module may embed a secret key that enables the system to determine the embedding location of the watermark in the host data. The second operating module, the watermark decoder, takes the watermarked host data as input and processes them to extract the watermark. If a secret embedding key was used during the encoding module, then the same key is required for the decoding process. Depending on the algorithm employed, the original host data may also be used explicitly to extract the watermark. The final module, authentication, compares the recovered watermark with the original watermark to estimate the similarity between the two sets. If the watermark is biometric in nature, then this biometric data can be optionally used by the authentication system.

Watermark encoding and decoding techniques fall into two categories: spatial and transform domain

Iris Digital Watermarking. **Table 2** Characteristics of biometric watermarking systems

Characteristic	Description
Imperceptibility	The degree to which the host image is visibly altered or distorted due to the presence of the watermark. Watermarked images that bear no visible difference from their original host image are said to contain imperceptible watermarks. Rarely, this characteristic may be evaluated beyond the scope of human perception. In these cases, a question is raised as to whether or not it is possible to reveal that an image contains a watermark through the aid of a machine or program (without access to the original host image)
Robustness	The ability of the watermark to be detected and extracted after the watermarked image has been subjected to any variety of transformations (i.e., compression, filters, affine transformations, etc.)
Fragility	The ability to detect any file transformation by way of the watermark. The detection might result from an inability to extract the watermark or from an extracted watermark that in not intact
Capacity	The amount of information that can be embedded in the host data of a watermarking system. This is a function of the type and size of the host data that are being watermarked and the robustness of the watermarking system in terms of detectability and extractability
Performability	The degree to which the watermark affects the performance of the biometric system(s) in question. At a minimum, biometric watermarks should not have an adverse affect on the performance of the biometric system(s) that they protect. Here, performance can entail matching error rates, image quality, efficiency of computation time, etc. Some biometric watermarking schemes may positively affect the performance of the biometric system



Iris Digital Watermarking. **Figure 2** Generalized block diagram of an iris digital watermarking process. Shaded blocks indicate the main watermarking modules, while dashed blocks/lines indicate optional areas of processing that are algorithm specific.

techniques. Each category has specific advantages and disadvantages, but in general, spatial domain techniques have lower complexity and offer higher robustness to biometric replacement attacks in which the host biometric region is replaced by an imposter's biometric data. On the other hand, transform domain techniques are of higher complexity, but are more robust to geometrical attacks such as rotation, scaling, and translation.

A number of studies have investigated different approaches to biometric watermarking. Ratha et al. [5] propose an algorithm for biometric watermarking to counter replay attacks in on-line fingerprint authentication systems. The authors modify the least significant bit (LSB) of the indices obtained as a result of applying wavelet scale quantization (WSQ) compression. The indices are chosen based on an embedding key that is used as a seed for a random number generator. Similarly, Noore et al. [6], utilize the Discrete Wavelet Transform (DWT) to watermark fingerprint images with face and demographic text data. In [7], Low et al. watermark a nonbiometric host image with off-line handwritten signature in the form of a discretized bit string. They experiment with three watermarking techniques: LSB, Code Division Multiple Access (CDMA) spread spectrum in the spatial domain, and CMDA spread spectrum in a transform domain such as the DWT. Their experiments show that CMDA in the wavelet domain provides the most robust results with respect to jpeg compression and image quality. In [8], a Quantization Index Modulation (QIM) watermarking technique is utilized to encode dynamic and static handwriting signature features into the host signature from which the features are extracted. Embedding locations are chosen based on the analysis of the signature in two transform domains: Ridglet and Radon-DCT (Discrete Cosine Transform). The authors conclude that although the static and dynamic features by themselves provide modest levels of security, fusion of both feature types improves security and performance.

Watermark Encoding and Decoding

The encoding portion of an iris digital watermarking scheme involves embedding the watermark in the original host data. The decoding portion, on the other hand,involves extracting the watermark from the host image. Bartlow et al. [9] present a watermarking technique modified from [10], based on amplitude modulation in the spatial domain to protect iris biometric systems. In the work, iris images are watermarked with simulated voice feature vectors. Equations 1, 2, and 3 correspond to the amplitude modulation scheme presented by Kutter in [10]. Equation 4 represents the adaptation required to watermark iris biometric data as presented in [9].

- **Encoding**- Amplitude modulation encodes the bits of the watermark bymodifying pixel intensities in images. Pixel intensities are modified bychanging values B_{ij}, in the blue channel of the RGB spectrum. These modifications occurmultiple times over the extent of the image and are either additive orsubtractive, depending on the value of the bit, *s*,and its proportionality to the luminance, L_{ij}, as seen in Equation 1.

$$B_{ij} \leftarrow B_{ij} + (2s-1)L_{ij}q \begin{cases} B_{ij} \in \text{embedding locations} \\ s \in \text{bit} \\ L_{ij} \in \text{Luminance} \\ q \in \text{Encoding strength} \end{cases} \quad (1)$$

- **Decoding**- The decoding process estimates the pixel value in the "encoded"image by considering a linear combination of the pixels in a cross-shaped neighborhood around the encoded bit as seen in Equation 2.

$$B_{ij} = \frac{1}{4c}\Big(\sum_{k=-c}^{c} B_{i+k,j} + \sum_{k=-c}^{c} B_{i,j+k} - 2B_{ij}\Big)\{c \in \textit{neighbor size} \quad (2)$$

After decoding and arriving at an estimated pixel B_{ij} value, the difference between the estimated and watermarked pixelsis averaged over all embedding locations for that bit. Finally, thesign of this value indicates the bit (if positive = 1, if negative = 0). However, to attenuate robustness to compression, cropping, and affine transformations, an adaptive thresholding method is introduced: two bits, 0 and 1, are appended to every bit stream as seen in Equation 3.

$$bit = \begin{cases} 1, \delta^b > \dfrac{\delta^0 + \delta^1}{2} \\ 0, otherwise. \end{cases} \begin{cases} \delta^0 \textit{ average diff of all 0 reference bits} \\ \delta^1 \textit{ average diff of all 1 reference bits} \\ \delta^b \textit{ average diff of current bit} \end{cases} \quad (3)$$

- **Adaptation to Iris Biometric Data**- Iris images for use in biometric systems are usually captured in the grayscale format. The encoding process has to be modified to take this into consideration. For example, [11] modifies the encoding equation to take in local image information such as gradient, P_{GM}, and standard deviation, P_{SD}, of the cross-shaped neighborhood to adjust the watermarking strength. Parameters A and B aid in adjusting the strength of the standard deviation and gradient while modulating the bits to be encoded (for all experiments $A = 100$, $B = 1{,}000$). The following equation represents this adaptation.

$$P_{WM}(i,j) = P(i,j) + (2s-1)P_{AV}(i,j)q \left(1 + \frac{P_{SD}(i,j)}{A}\right)\left(1 + \frac{P_{GM}(i,j)}{B}\right) \quad (4)$$

P_{AV} represents the average pixels in a 5 × 5 cross-shaped neighborhood centered around i, j. Finally, the reconstructed image can be calculated by replacing the watermarked bit with the pixel value obtained from Equation 2.

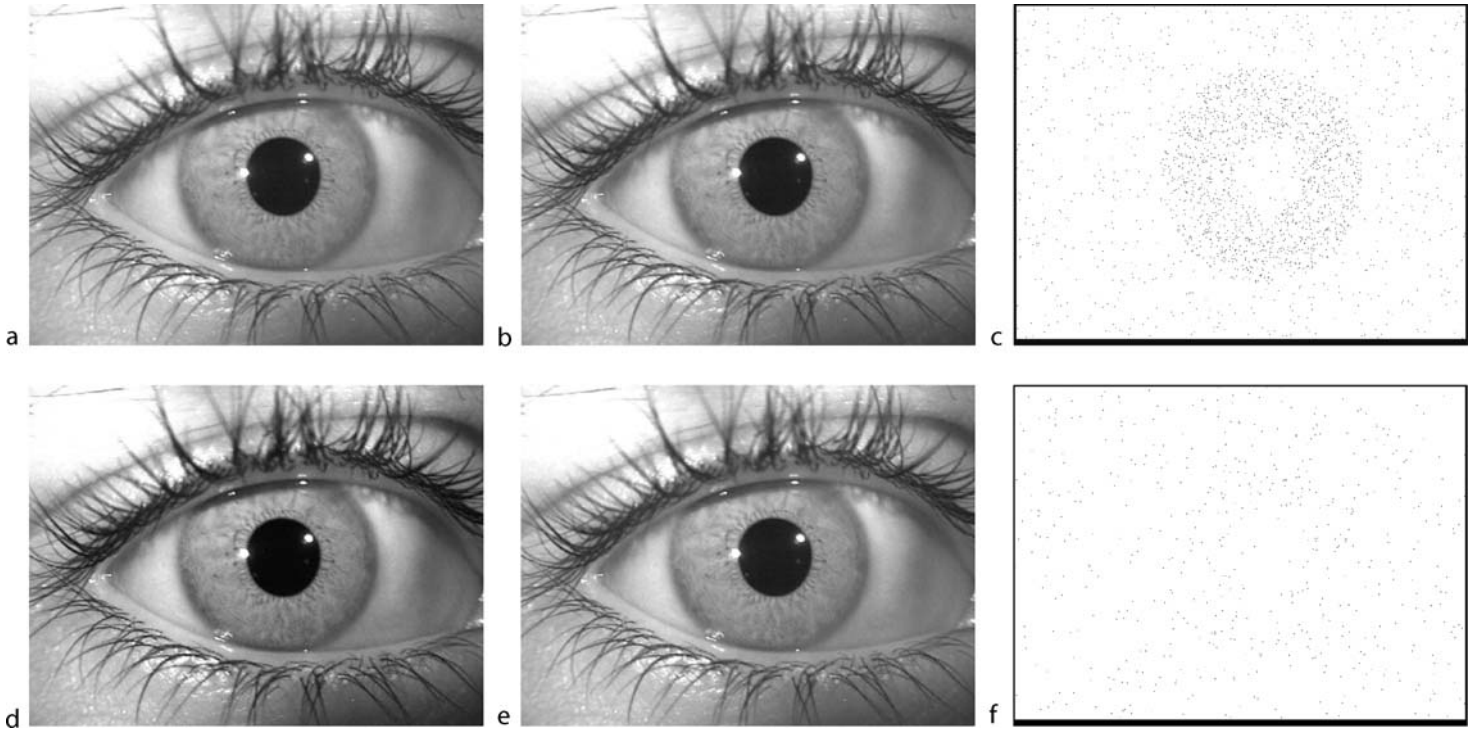
A crucial issue while encoding the watermark in the host image relates to the degree to which it affects the matching performance of the host biometric. The algorithm in [9] further extends [11] by including a parameter that indicates the proportion of the watermark encoded in the iris region of the image. Figure 3 visualizes the variation of this parameter by examining two extreme cases: the top row shows the most perceptible case and the bottom row shows the least perceptible case tested in the work. This location-specific encoding option may be beneficial not only to iris digital watermarking but also to biometric watermarking in general.

Application Scenarios and Attacks

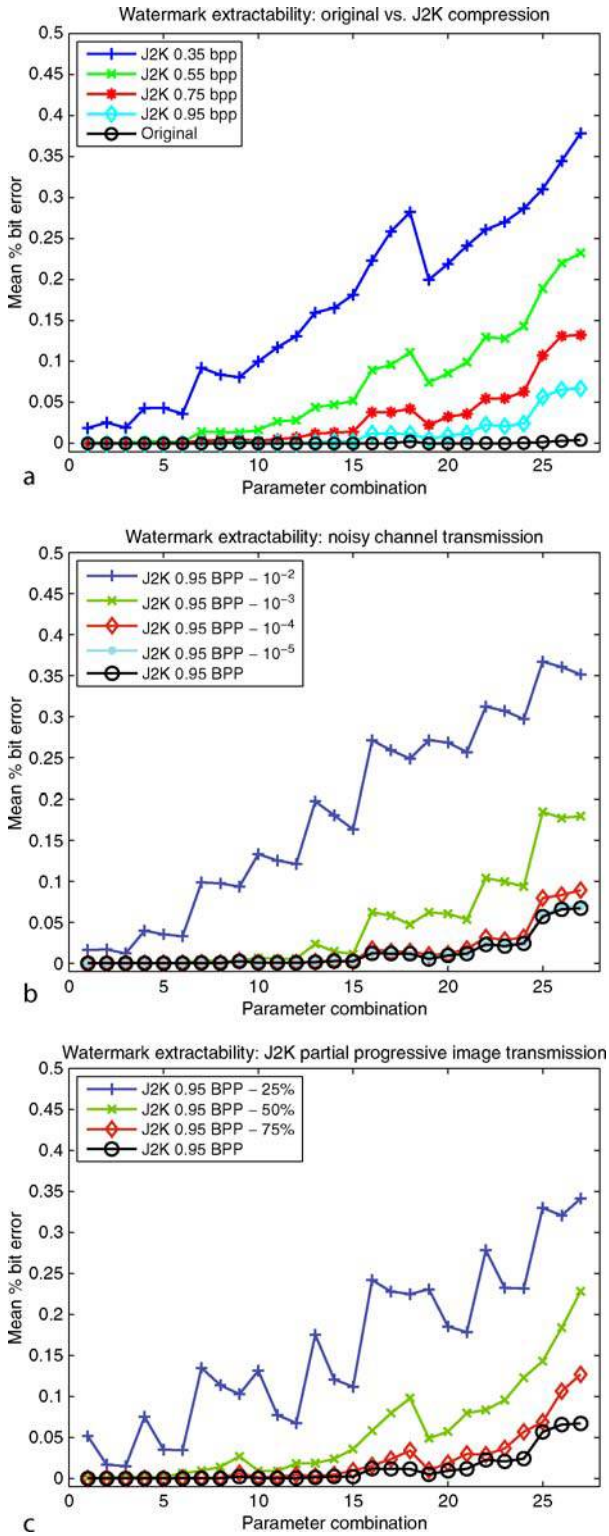
Depending on the intended use, iris digital watermarking systems are vulnerable to a series of application scenarios and attacks. Application scenarios can be thought of as normal usage patterns that a watermarking system should realistically be expected to withstand without serious side effects on the performance of any of the characteristics outlined in Table 2. Examples of application scenarios can include, but are not limited to database (re)compression, partial progressive decoding, and noisy channel transmission. Each of these scenarios can have an effect on one or more characteristics of a biometric watermarking system. For instance, a highly compressed watermarked image may lead to difficulties in the watermark extraction process, as a compression algorithm often significantly alters an image, which in turn alters the watermark itself. Occasionally, operational environments supporting slow data transmission speeds may force a system to progressively decode portions of an image as it becomes available. This type of application scenario can have an effect on the robustness of the extraction process, the performance of the biometric system(s) in question, and potentially the imperceptibility of the watermark. In [9], the authors study the robustness of the decoder after the application of these three scenarios. Figure 4 shows the performance of the decoder as measured by the mean percentage bit error considering 27 different parameter combinations of the watermarking system (parameters are specifically outlined in Table 3).

The graphs demonstrate the ability of the technique to tolerate the application scenarios, and in many cases, with little or no effect on the extraction process. Although this is just one example, many techniques such as those described in [3] are capable of tolerating scenarios similar to these. However, the expected application environment should be considered while selecting a specific watermarking technique, since a given technique may be well-suited for one application scenario but ill-suited for another [3].

Perhaps the most notable difference between the general field of digital watermarking and biometric watermarking is its relationship to the measure of performability. For obvious reasons, a biometric watermarking system must minimize the effect it has on the biometric system it protects. Issues such as matching performance, image quality, computational efficiency, and even legal repercussions must not be ignored. A biometric watermarking system should not negatively impact on the main modules of the biometric system(s) in question, viz., the feature extraction and matching modules. It should be noted that this effect could potentially propagate itself in two ways. Perhaps, the most obvious effect is when the host data are used in a biometric system; here, the presence of the watermark may impede the feature extraction process by adding noise to the image. Naturally this can lead to inaccuracies in the matching module. A less obvious effect is



Iris Digital Watermarking. **Figure 3** Perceptability of Watermarked Images. Encodings refers to the number of times a watermark is embedded into the host image.



Iris Digital Watermarking. Figure 4 Robustness of the decoder in three application scenarios: (**a**) Varying levels of J2K compression (measured in bits per pixel (bpp); (**b**) J2K compression + varying levels of zero mean white Gaussian noise (variances shown in legend); and (**c**) J2K compression + varying levels of partial progressive decoding (% decoded shown in legend).

Iris Digital Watermarking. Table 3 Watermarking parameter combinations examined in [9]. The first item of the 3-tuple represents the degree of pixel modulation; second item represents the number of times the watermark is embedded in the image; the third item represents the proportion of the watermark embedded in the iris portion of the host image. Further explanations can be found in [9]

1	0.10-60-0.67	**2**	0.10-60-0.33	**3**	0.10-60-0.00	**4**	0.10-40-0.67	**5**	0.10-40-0.33	**6**	0.10-40-0.00	**7**	0.10-20-0.67	**8**	0.10-20-0.33	**9**	0.10-20-0.00
10	0.06-60-0.67	**11**	0.06-60-0.33	**12**	0.06-60-0.00	**13**	0.06-40-0.67	**14**	0.06-40-0.33	**15**	0.06-40-0.00	**16**	0.06-20-0.67	**17**	0.06-20-0.33	**18**	0.06-20-0.00
19	0.04-60-0.67	**20**	0.04-60-0.33	**21**	0.04-60-0.00	**22**	0.04-40-0.67	**23**	0.04-40-0.33	**24**	0.04-40-0.00	**25**	0.04-20-0.67	**26**	0.04-20-0.33	**27**	0.04-20-0.00

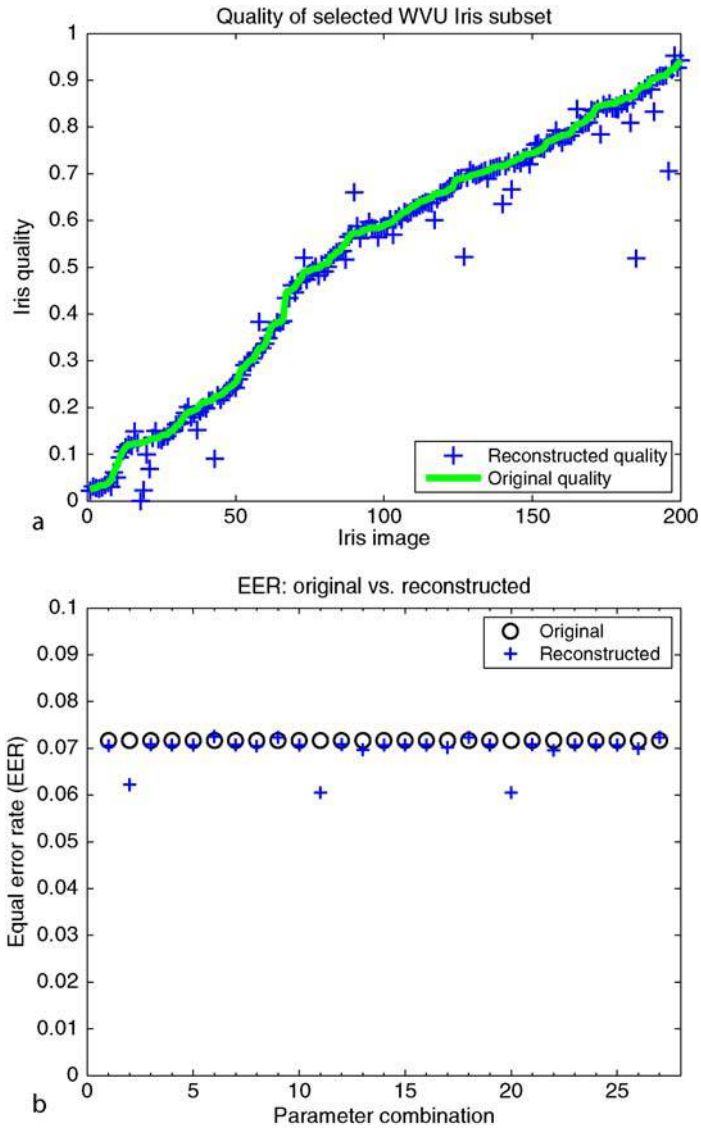
when a biometric feature vector serves as a watermark and is also used in the authentication stage. In this scenario, accurate extraction of the watermark is of utmost importance as small changes in the values of feature vectors can lead to significant changes in authentication results. Although little or no work exists studying the latter of the two effects, we once again refer to [9] where the effect of an amplitude modulation watermarking system on biometric image quality and matching performance is studied. Figure 5 shows the effect on image quality and matching performance of the underlying iris biometric system. The graphs compare the quality and matching results from the original host images with the images resulting after watermark extraction and image reconstruction, and the tables provide averaged results across the entire dataset.

Much like the three application scenarios mentioned earlier, these results provide one example of a biometric watermarking system that does not produce unwanted side effects. In particular, image quality is seldom affected with any level of significance and matching rates also do not suffer from performance degradation.

Application scenarios aside, biometric watermarking systems must also deal with attacks or malicious attempts to subvert a system. Such attacks may involve removal, alteration, or replacement of the embedded watermark found in an image. Although some of these may fall in the application scenarios described earlier, examples of attacks include rotation, scaling, translation, cropping, masking, and (re)watermarking. As in the case of application scenarios, different biometric watermarking techniques can handle different attacks with varying degrees of success. Although not specific to biometric watermarking, Zheng et al. provide an excellent description of so-called RST (rotation, scaling, translation) invariant watermarking algorithms in [3]. Often the ability to handle a given attack lies in the domain in which a biometric watermarking technique operates. For instance, rotation attacks are handled with greater ease by watermarking techniques that operate in transform domains (i.e., Fourier, DCT, wavelet, etc.). This type of attack is arguably more difficult to handle by techniques that operate in the spatial domain. Conversely, techniques in the spatial domain are more likely to handle biometric replacement attacks (i.e., replacing the watermarked iris or face region of an image), as the biometric ROI can be easily located in the spatial domain, but more difficult to localize in a transform domain.

Patents, Tools, and Commercial Products

Searching the United States Patent Office (USPO) for "Biometric Watermarking" yields well over 100 entries of varying relevance to the field. The most notable is titled "Biometric Watermarks" and was issued in 2001 to GTE Service Corporation [12]. This patent outlines the general schematic for a biometric watermarking system. A more recent patent issued to Canon by the USPO relates specifically to iris digital watermarking and includes an embedded watermarking system in a camera with the intended purpose of associating a photographers biometric information with images taken by the camera [13]. There are a few freely available tools related to digital watermarking and biometric watermarking exist. For example, Stirmark



Iris Digital Watermarking. **Figure 5** Effect of watermarking on biometric performability: (**a**) Compares the quality of 200 original host data images vs. the reconstructed images after watermark extraction; (**b**) Compares the EER across 100 users before watermarking and after image reconstruction; (**c**) Average image quality before watermarking and after reconstruction; (**d**) Average EER (%) before watermarking and after reconstruction across 100 subjects.

Benchmark 4.0 is a software tool designed to perform robustness testing of image watermarking algorithms [14]. Another tool, Checkmark, also provides a bed of attacks to evaluate the robustness of a watermarking system [15]. Many commercial entities offer a broad range of digital watermarking solutions that can potentially fall under the category of biometric watermarking. Perhaps the most widely known of such companies is DigiMarc Corporation based in Oregon, US.

Summary

Iris digital watermarking is a technique utilized in tandem with cryptographic systems to protect iris biometric images. The watermarking scheme can be used as a mechanism for proving file authenticity, tracking chain of custody and data reproduction, or affording a multimodal biometric option, thereby offering an additional layer of security after data decryption. A wide range of watermarking systems exist that

operate in any one of several domains (spatial, Fourier, DCT, wavelet, etc.). A system's expected application profile and threat model will dictate the choice of watermarking algorithm, the nature of the watermark to utilize, and a viable set of algorithmic parameters. Carefully making these decisions will result in a formidable layer of postdecryption protection without compromising on the performance aspects of the underlying biometric system(s).

Related Entries

- ▶ Binding of Biometric and User Data
- ▶ Biometric Encryption
- ▶ Biometric System Design

References

1. Tirkel, A.Z., Rankin, G.A., van Schyndel, R.M., Ho, W.J., Mee, N.R.A., Osborne, C.F.: Electronic Watermark pp. 666–673 (1993)
2. Artz, D.: Digital steganography: hiding data within data. Internet Computing, IEEE **5**(3), 75–80 (May/Jun 2001). DOI 10.1109/4236.935180
3. Zheng, D., Liu, Y., Zhao, J., El-Saddik, A.: A Survey of RST Invariant Image Watermarking Algorithms. ACM Comput. Surv. **39**(2), (2007)
4. Ahmad, S., Lu, Z.M.: A joint biometrics and watermarking based framework for fingerprinting, copyright protection, proof of ownership, and security applications. Computational Intelligence and Security Workshops, 2007. CISW 2007. International Conference on pp. 676–679 (15-19 Dec. 2007)
5. Ratha, N.K., Connell, J.H., Bolle, R.M.: Secure data hiding in wavelet compressed fingerprint images. In: MULTIMEDIA '00: Proceedings of the 2000 ACM workshops on Multimedia, pp. 127–130. ACM, New York, NY, USA (2000). DOI http://doi.acm.org/10.1145/357744.357902
6. Noore, A., Singh, R., Vatsa, M., Houck, M.M.: Enhancing security of fingerprints through contextual biometric watermarking. Forensic Science International **169**, 188–194
7. Low, C.Y., Teoh, A.B.J., Tee, C.: A preliminary study on biometric watermarking for offline handwritten signature. Telecommunications and Malaysia International Conference on Communications, 2007. ICT-MICC 2007. IEEE International Conference on pp. 691–696 (14-17 May 2007)
8. Maiorana, E., Campisi, P., Neri, A.: Biometric signature authentication using radon transform-based watermarking techniques. Biometrics Symposium, 2007 pp. 1–6 (11-13 Sept. 2007). DOI 10.1109/BCC.2007.4430543
9. Bartlow, N., Kalka, N., Cukic, B., Ross, A.: Protecting iris images through asymmetric digital watermarking. Automatic Identification Advanced Technologies, 2007 IEEE Workshop on pp. 192–197 (7-8 June 2007). DOI 10.1109/AUTOID.2007.380618
10. Kutter, M., Jordan, F., Bossen, F.: Digital signature of color images using amplitude modulation. In: Proc. SPIE EI, San Jose, CA, pp. 518–526 (1997)
11. Jain, A.K., Uludag, U., Hsu, R.L.: Hiding a face in a fingerprint image. In: ICPR '02: Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 3, p. 30756. IEEE Computer Society, Washington, DC, USA (2002)
12. Musgrave, C.: Biometric watermarks. Patent 6,208,746, U.S. Patent and Trademark Office, Washington, D.C. (2001)
13. Morikawa, G., Tokura, G.: Picture taking apparatus and method of controlling same. Patent 7,305,089, U.S. Patent and Trademark Office, Washington, D.C. (2007)
14. Petitcolas, F.A., Anderson, R.J., Kuhn, M.G.: Attacks on copyright marking systems. In: Information Hiding, pp. 218–238 (1998). citeseer.ist.psu.edu/petitcolas98attacks.html
15. Pereira, S., Voloshynovskiy, S., Madueno, M., Marchand-Maillet, S., Pun, T.: Second generation benchmarking and application oriented evaluation In: IHW '01: Proceedings of the 4th International Workshop on Information Hiding, pp. 340–353 Springer, London, UK (2001)

I

Iris Encoding and Recognition using Gabor Wavelets

John Daugman
Cambridge University, Cambridge, UK

Synonyms

Daugman algorithm; IrisCode; Iris2pi

Definition

The method of encoding iris patterns that is used in all current public deployments of iris recognition technology is based on a set of mathematical functions called ▶ Gabor wavelets that analyze and extract the unique texture of an iris. They encode it in terms of its phase structure at multiple scales of analysis. When this phase information is coarsely quantized, it creates a random bit stream that is sufficiently stable for a given eye, yet random and diverse for different eyes, that iris patterns can be recognized very rapidly and reliably over large databases by a simple test of statistical independence. The success of this biometric algorithm may be attributed in part to certain important properties of the Gabor wavelets as encoders, and to the simplicity

and efficiency of searches for matches when pattern information is represented in terms of such phase bit strings.

Introduction

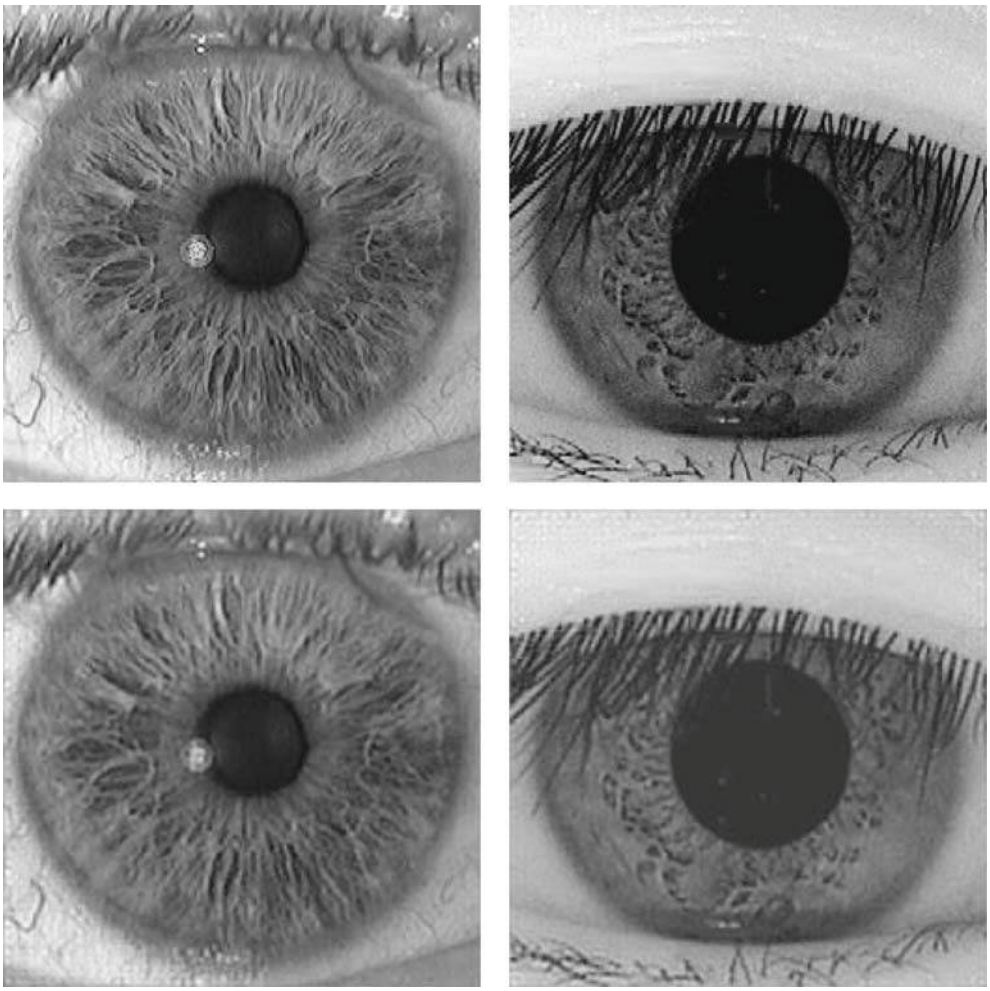
Different biometric modalities use diverse methods for encoding the features on which they depend. The overall goal in designing biometric encoders is always the same – maximizing between-class variation while minimizing within-class variation – but very different strategies have been developed for representing the chosen features and their random variation. Even within a single modality, such as fingerprint recognition, some methods compile lists of discrete minutiae coordinates and angles, while other methods encode global ridge-flow descriptions. Facial representations may be two-dimensional ("appearance-based") or three-dimensional "model-based"); may try to achieve some degree of pose-invariance, illumination-invariance, or expression-invariance; and the scale of analysis may be global (e.g., eigenface decompositions) or focused on local, high-resolution, detail (e.g., skin texture analysis). In the case of iris patterns, one does not find any easily enumerated lists of distinct features like the fingerprint minutiae, nor indeed any sets of features that even possess established names. Rather, one finds a plethora of textures spanning many scales of analysis, a wide spatial frequency range, and which might be described using many candidate image statistics. Whereas a natural feature to mark in a fingerprint is a ridge ending or a bifurcation, in the case of iris patterns we find random features with no simple geometric or graph-like structure, that adhere to no taxonomy, and that are defined across many different size ranges. We need a language rich enough to capture subtleties like "mottling" and "modulation," yet simple enough that all instances are commensurable, and powerful enough to deliver extremely rapid and confident recognition decisions. What might be such a language?

Most naturally occurring textures lend themselves well to mathematical description in terms of both spectral (Fourier-based) and spatially localized properties. The spectral language captures predominant undulations and quasi-coherences that are the essence of texture, whose interwoven appearance reveals why this word shares a root with "textile:" *texere* – "to weave." But at the same time, spatial variation in the undulations destroys any simple coherence, and breaks symmetries. To describe natural textures effectively and efficiently, for example in order to build a biometric recognition system based on iris textures, we need a language that is able to specify *both* the spatial and the spectral properties well.

The upper panels of Fig. 1 illustrate two human iris patterns, Caucasian on the left, oriental on the right. They reveal rich and unique textural structure, and also characteristic ethnic *differences* in their appearance. For example, the Caucasian iris has finer detail and longer radial correlations, whereas the oriental iris has somewhat coarser and more isotropic (less elongated) features, which are also more concentrated near the pupil. All of these general aspects of iris patterns – their undulatory textures, their spatial variation, their statistical structure in correlation distances, and, above all, their rich randomness, motivated the author nearly two decades ago to develop a method for encoding and recognizing iris patterns using a particular mathematical family of localized undulations called ▶ Gabor wavelets. Today, that algorithm [1, 2] remains the method used in all public deployments of iris recognition. This chapter reviews some of the essential aspects of this approach to image analysis, encoding, and pattern recognition.

Gabor Wavelets as a Complete Image Basis

Classical signal processing divided broadly into analysis performed in the *signal domain* (time in the case of time-varying signals like sound waveforms, or space in the case of images), versus analysis performed in the *Fourier domain* in which the signal is represented as a linear combination, or superposition, of global sinusoids. Although a bedrock of signal processing, Fourier analysis is hampered by its excessively global perspective: in the Fourier domain, every coefficient associated with a particular frequency component summarizes the presence of that frequency component over the entire extent of the signal. Similarly, every local point in the signal has an influence on every coefficient in the Fourier domain (i.e., on every frequency component). This extreme, reciprocal, mapping "from local to global" and "from global to local" across the two domains is reflected in the fact that their respective independent variables are themselves reciprocals: time versus



Iris Encoding and Recognition using Gabor Wavelets. **Figure 1** The upper two panels illustrate the rich textures found in the iris, and also some typical ethnic differences: on the left is a Caucasian iris, and on the right an oriental iris. Some significant differences include the longer radial correlation distances and finer detail in (especially blue-eyed) Caucasians, and the tendency of oriental eyes to have more of their iris texture near the pupil. The lower two panels show reconstructions of the upper two images using only a sparse discrete set of 2D Gabor wavelets, incorporating just six spatial frequencies (one octave apart) and six orientations, as seen in Fig. 2.

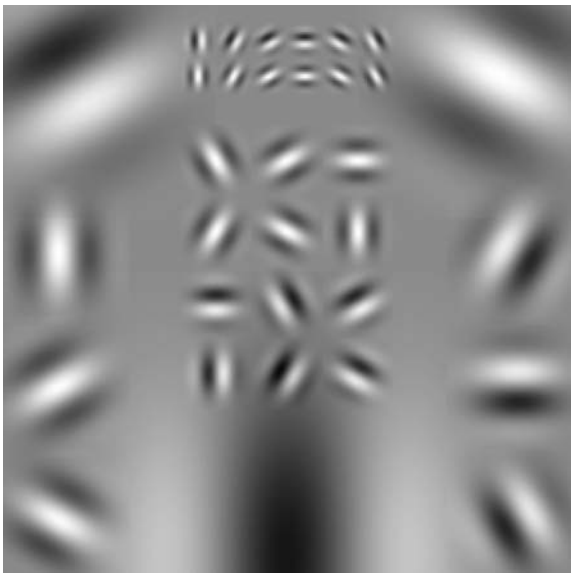
(temporal) frequency, as Hertz (1/s); and for images, space versus spatial frequency (cycles per degree).

In the 1980s a number of mathematicians, most of them Francophone, began to formulate a new synthesis of these two domains based on a kind of compromise between the global (spectral) and local (punctate) perspectives. It came to be called *wavelet theory* (translated from a French neologism "*ondelette*" – a small wave). The key insight of wavelet theorists like Yves Meyer [3] and Ingrid Daubechies [4] was that it was possible to construct complete representations of functions by superposition of a set of universal elementary functions all of which were dilates and translates of a single shape on finite support, forming a *dyadic set*. Typically the successive dilation (stretching) factors were powers of 2, and the translation intervals scaled reciprocally, so that *self-similarity* was preserved across all scales. Originally there were five requirements for such sets of basis functions to be deemed wavelets: they must all be (1) dilates and (2) translates of each other; they must have (3) strictly compact support; (4) all of their derivatives must exist; and (5) they must be mutually orthogonal (i.e., have zero inner products with each other). These constraints and admissibility

conditions have varied over the years. But perhaps the main legacy of this development of what is also called *multiresolution analysis* was the unification of the local and the spectral perspectives into a single framework: as the name "wavelet" implies, the terms of analysis became simultaneously local, yet frequency-specific. One important practical manifestation of these theoretical developments is found in JPEG-2000 compressive image encoding, which is based on a class of wavelets developed by Ingrid Daubechies [4].

But as early as 1946 a class of wavelet-like elementary functions named *logons* (from the Greek word for "order") had been proposed by the Hungarian physicist Denis Gabor [5], who also invented holography and won the Nobel Prize in 1971. Although these functions lack some of the stricter requirements on wavelets, they satisfy more lax definitions and they have certain advantages including being expressible in closed analytic form (i.e., they can be defined in terms of classical functions). They take the form of complex exponentials (i.e., Fourier components) multiplied by Gaussian envelopes, which localize them and specify their scale. Whether or not their parameters are constrained to maintain self-similar profiles at all scales (a configuration not anticipated by Gabor), these wavelets can form a "frame" and can be used as a basis for complete expansions: any signal or image can be constructed as a linear combination, or superposition, of such wavelets. This is illustrated in the lower two panels of Fig. 1: those two reconstructions were synthesized entirely from a discrete dyadic set of self-similar complex Gabor wavelets using only six frequencies and six orientations, as shown in Fig. 2.

The library of self-similar Gabor wavelets whose real and imaginary parts are portrayed pictorially in Fig. 2 differ from each other in frequency by steps of one octave (successively doubling in frequency, while their Gaussian widths are successively halved). They are defined in each of six orientations. Being complex functions, their parts have two phases: cosine (even-symmetric) and sine (odd-symmetric). This library emulates the architecture found in the brain's visual cortex, whose neural receptive fields are structured for sequential orientation selectivity [6], size or spatial frequency selectivity with roughly one octave half-bandwidth and receptive field profiles [7] whose excitatory/inhibitory inputs resemble the structures seen in Fig. 2 with quadrature (90°) phase-tuned elements arranged in sine/cosine pairs [8]. Taken



Iris Encoding and Recognition using Gabor Wavelets.

Figure 2 Visual library of real and imaginary parts of the 2D Gabor wavelets, defined in six discrete orientations and in six discrete frequencies that differ from each other in one octave steps (i.e., by successive factors of two). The lowest of the six frequencies is not included here as it fills the entire image. This discrete set of wavelets is the set that was used to synthesize the two iris images as shown in the lower two panels of Fig. 1, reconstructing the original natural images seen in the upper two panels.

together, these empirical neurophysiological observations support the "2D Gabor model" [9] of image representation in the brain's visual cortex, but practical engineering implementation of it is complicated by the fact that these wavelets constitute a nonorthogonal set. Their lack of mutual independence (their nonzero inner product) has the consequence that the coefficients needed for image expansion or reconstruction are not the same as the coefficients obtained simply by projecting the image onto the wavelets; they cannot be obtained merely by filtering or convolution with the image. A solution for finding correct expansion coefficients so that the wavelets can be used as a complete image basis is a "relaxation network" [10]. This method is how the synthetic iris images in the lower panels of Fig. 1 were constructed, using only the discrete set of wavelets seen in Fig. 2 having six frequencies, one octave apart (the lowest frequency wavelet being omitted as it fills the entire image) and six orientations. It is

clear that the superposition, or linear combination, of these discrete wavelets using appropriately computed coefficients converges faithfully to the original images in the upper panels, up to the resolution determined by the highest frequency wavelet used. Thus the discrete ensemble of computed wavelet coefficients, which may be called a complete discrete 2D ▶ Gabor Transform [10], capture all the information in the original images and, more importantly, constitute an extremely useful representation of it.

Gabor Wavelets and the Uncertainty Principle

The evident richness of natural iris textures, as illustrated in Fig. 1, invites description that is specific both in spectral terms (the frequencies and orientations of interwoven undulations), and in spatially localized terms. Yet these two goals are in mutual conflict, because of a fundamental *Uncertainty Principle* [5, 9] that makes the resolution of either type of information possible only at the expense of resolution for the other. The Uncertainty Principle is a fundamental law of mathematics, not simply an empirical problem; it can be derived as a general relationship constraining functions and their Fourier transforms. One particular instantiation of it is the familiar Heisenberg Uncertainty Principle in quantum physics: the position and the momentum of a particle cannot be known with simultaneously unlimited accuracy, given that its momentum is interpretable as wavelength and therefore has spectral specificity. The abstract form of the Uncertainty Principle asserts a lower bound on the product of the "effective width" of any function and that of its Fourier transform. The functions that uniquely achieve this lower bound, and therefore achieve maximal specificity or localizability in both domains at once, are the (complex-valued) Gabor wavelets [5].

Defined in two dimensions with (x,y) interpretable as image coordinates, these wavelets have the following parameterized functional form [9]:

$$f(x,y) = e^{-\left[(x-x_0)^2/\alpha^2+(y-y_0)^2/\beta^2\right]} e^{i\left[u_0(x-x_0)+v_0(y-y_0)\right]}, \tag{1}$$

where (x_0,y_0) specify the wavelet's center position in the image, (α,β) specify its effective width and length, and (u_0,v_0) specify its modulation, which has spatial frequency $\omega_0 = \sqrt{u_0^2+v_0^2}$ and orientation $\theta_0 = \arctan(v_0/u_0)$. (A further degree-of-freedom not included above is the relative orientation of the elliptic Gaussian envelope, which creates cross-terms in xy.) The 2D Fourier transform $F(u,v)$ of a 2D Gabor wavelet has exactly the same functional form, with parameters just interchanged or inverted:

$$F(u,v) = e^{-\left[(u-u_0)^2\alpha^2+(v-v_0)^2\beta^2\right]} e^{-i\left[x_0(u-u_0)+y_0(v-v_0)\right]} \tag{2}$$

Thus Gabor wavelets are *self-Fourier*, since $f(x,y)$ has the same form as $F(u,v)$. The modulation parameters (u_0,v_0) in the image domain play the role of location parameters in the Fourier domain, specifying a wavelet's peak frequency and orientation sensitivity if used as a filter. The width and length parameters α and β which set the effective size of the Gaussian envelopes play reciprocal roles in the two domains: the larger a wavelet is in one domain, the smaller it is in the other, as dictated by the Uncertainty Principle. Finally, some further interesting properties of Gabor wavelets besides their completeness (ability to be an expansion basis for other functions, like the iris images in Fig. 1) and their self-Fourier property, are that as a family of functions they are closed under multiplication and under convolution: the product of any two Gabor wavelets is just another Gabor wavelet; and indeed the convolution of any two Gabor wavelets is also just another Gabor wavelet. For our present purposes, their most useful property besides their optimal joint specificity in both spatial and spectral terms is their utility for *analyzing* image structure, including defining the *phase* of any element of an image since the wavelets are complex-valued, and the utility of such descriptions for pattern recognition.

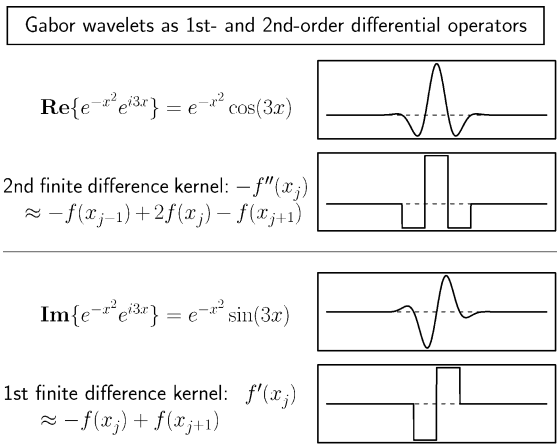
Gabor Wavelets and the Calculus

It is clear from the functional form defining $f(x,y)$ above that Gabor wavelets reduce to pure Fourier components when the Gaussian space constants (α,β) become large. Then the use of these functions for image analysis becomes equivalent to Fourier analysis: the Fourier transform is just a special case of a Gabor transform. At the other extreme, in the limit that (α,β) become small, the functions reduce to delta functions,

simply sampling particular points in the image. Thus, the Gaussian scale parameter essentially creates a continuum that bridges the dichotomy between local (point sampling) and global (Fourier) analysis, embracing those classical approaches as the endpoints of the continuum. At points between those two extremes, the wavelets enable a kind of local spectral analysis to be performed, extracting Fourier-like information (e.g., phase and frequency descriptions) but in a local region-specific fashion.

When Gabor wavelets are used as filters to convolve with a signal, their effect depends of course on the values chosen for their parameters. If the size of the Gaussian is large compared with the modulation wavelength, allowing several cycles of oscillation before attenuation, then the complex wavelet becomes a narrowband filter that allows a well-defined phase to be assigned to each point in the output signal. Specifically, the phase assigned to a point is the arctangent of the ratio of the imaginary part to the real part of the complex-valued result of the convolution with the signal at that point. But for smaller Gaussian space constants that allow only one or two cycles of oscillation before attenuation, the wavelets behave instead like approximate first- and second-order differential operators.

Figure 3 illustrates (for the one-dimensional case) how convolution with such wavelets approximates taking the first or second derivative of a signal. The real and imaginary parts of a Gabor wavelet having such a parameterisation are plotted in the first and third panels. The second panel plots the *second finite difference kernel,* which is the discrete filter that should be convolved with a discrete signal (a signal defined only on a discrete domain, such as the integers or a regular sampling lattice) in order to obtain the discrete approximation to a second derivative. There is an obvious resemblance between the continuous and the discrete functions plotted in the first and second panels. Likewise, the fourth panel plots the discrete approximation for a first-derivative operator, called the *first finite difference kernel,* which resembles the imaginary part of a Gabor wavelet (third panel). The definitions of the finite difference approximations provided on the left in this Figure correspond to Isaac Newton's [11] formulae for estimating the first or second derivatives ("*Fluxions*") of functions by combining adjacent sample values on regular unit sampling intervals, using weighting coefficients such as $[-1, +2, -1]$. In summary, the information extracted by convolving a signal with Gabor wavelets having parameterizations as indicated in Fig. 3 is closely related to the information obtained when the machinery of the Calculus is used to extract the first and second derivatives of a function. These properties contribute to the richness of the repertoire deployed by using complex-valued Gabor wavelets for image coding and analysis.



Iris Encoding and Recognition using Gabor Wavelets.

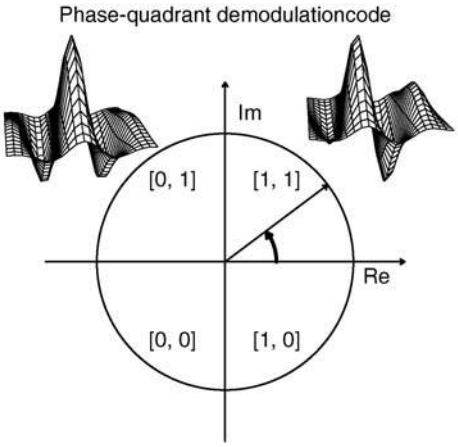
Figure 3 Analyzing and encoding signals or data using Gabor wavelets with narrow Gaussians corresponds to estimating a signal's first- and second-derivatives (or finite differences). These are approximated by convolving the signal with the imaginary and real parts, respectively, of a complex Gabor wavelet as shown in the right column. Such operations correspond simply to weighting adjacent samples algebraically (left column) with weights such as $[-1, +1]$ or $[-1, +2, -1]$ as noted in 1671 by Isaac Newton [11] in his theory of *Fluxions*.

Gabor Wavelets for Iris Recognition

The goals of pattern *recognition* are, of course, rather different from those of image encoding or analysis perse. However, when designing a pattern recognition system, it is nice to know that the image representation chosen is in principle complete, meaning that all information is available in the encoding as demonstrated by reconstructibility (Fig. 1), and also that interesting operations are implementable in the encoding such as extracting phase structure or derivatives (Fig. 3).

Interest in phase structure and in the zero-crossings of bandpass signals was invigorated three decades ago by some surprising proofs and critical demonstrations of complete signal reconstructibility from either type of information alone [12]. Phase and zero-crossings information are closely related: when a signal has been bandpass-filtered (so that it has zero mean) and digitized, then the most-significant-bit (MSB) of its samples corresponds to the most fundamental phase information, the *sign bit*; and of course this bit tracks the signal's zero-crossings. In order to conjoin such signal descriptions with representations on which decisions about pattern identity can be made, we need a kind of conceptual *signal-to-symbol converter*. One very efficient way to bridge this gap is to deploy the multi-scale Gabor wavelets as ▶ logico-linear operators, constructing bit streams from the quantization of the phase information that the wavelets extract. In the case of iris recognition, such a bit stream is called an ▶ IrisCode, and it allows identification decisions to be based on a simple test of statistical independence [1].

Figure 4 illustrates how the paired real and imaginary parts of 2D Gabor wavelets can be used to construct a *phase demodulation code*. Local patches of the image are projected onto both parts of the complex-valued wavelets, and each pair of resulting inner products constitute the real and imaginary parts of a complex number. Such a number has a phase and a modulus corresponding to its polar components in the complex plane, as portrayed in the phasor diagram. If one chose to resolve phase angles to an accuracy of only four quadrants as shown, then one would be extracting just two bits of phase information per wavelet. How finely should phase be quantized when encoded? How spatially fine in size (how high in frequency) should the discrete set of wavelets get? Both of these questions relate to the number of degrees-of-freedom one wishes to encode, and to how accurately one can realign encodings of subsequent images of the same iris in correspondence with an earlier image of it. Sharpening the finest scale of detail extracted makes the code for a given iris more detailed and more unique (thereby further decreasing the likelihood of False Matches), but by increasing the amount of minute detail that must be matched, such a strategy also increases the odds of failures-to-match (False non-Matches) due to uncertainties or inadequacy in registration and alignment.



$$h_{Re} = 1 \text{ if Re} \int_\rho \int_\phi e^{-i\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} I(\rho, \phi) \rho d\rho d\phi \geq 0$$

$$h_{Re} = 0 \text{ if Re} \int_\rho \int_\phi e^{-i\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} I(\rho, \phi) \rho d\rho d\phi < 0$$

$$h_{Im} = 1 \text{ if Im} \int_\rho \int_\phi e^{-i\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} I(\rho, \phi) \rho d\rho d\phi \geq 0$$

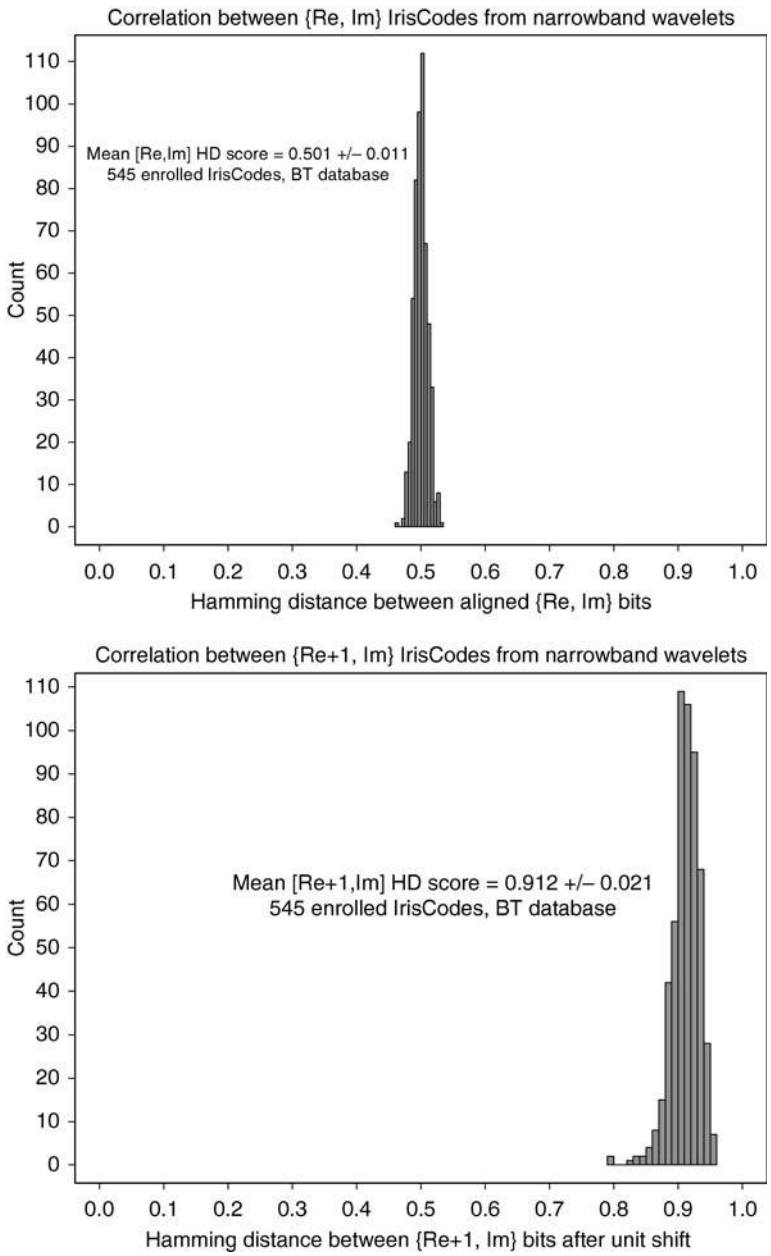
$$h_{Im} = 0 \text{ if Im} \int_\rho \int_\phi e^{-i\omega(\theta_0-\phi)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\phi)^2/\beta^2} I(\rho, \phi) \rho d\rho d\phi < 0$$

Iris Encoding and Recognition using Gabor Wavelets.
Figure 4 Constructing an IrisCode from phase analysis of iris texture. The two surfaces plotted are the real and imaginary parts of one 2D Gabor wavelet. Projecting a local area of the iris image onto these functions and integrating their products produces a complex number, whose real and imaginary parts specify a phasor in the complex plane as illustrated. Phase angle can be quantized at a chosen resolution accuracy (in this diagram, two bits for four quadrants) to create a phase-based IrisCode. The four equations give the projection integrals that specify the bits.

Code design issues also involve other aspects of the Gabor wavelet parameterization, including their *bandwidth*, which is determined by the effective number of oscillatory cycles contained within the Gaussian envelope before attenuation. More cycles cause narrower bandwidth: filters that are more sharply tuned. Related to this issue is the presence of a DC term (nonzero area or volume in the integral) in the real part of a Gabor wavelet if its bandwidth is broad (has only few cycles within the Gaussian); this is undesirable because it introduces code bit dependence on the overall brightness of the image, which ought to be irrelevant. However, the DC term can be nulled to zero when determining the sampling rate of the discrete taps which discretize the continuous wavelet, or alternatively by tiny adjustments in the discrete taps

themselves. A more fundamental issue when specifying code design is the phase coherence introduced by the wavelets if their bandwidth is narrow, because this reduces the randomness among the bits extracted by the code.

This issue is illustrated in Fig. 5, whose panels show the relationship between the real and imaginary parts of an IrisCode. Any bit of an IrisCode has equal a priori probabilities of being set or clear (ignoring the detection and masking of eyelids, eyelashes, reflections, or



Iris Encoding and Recognition using Gabor Wavelets. **Figure 5** Correlations between the real and imaginary parts of IrisCodes constructed from narrowband wavelets containing several cycles. Although orthogonality of the two quadrature pair components of a wavelet ensures that corresponding real and imaginary bit pairs are independent (upper panel), they show strong correlation when the two bit streams are simply shifted relative to each other by an amount corresponding to $\pi/2$ in wavelet phase (lower panel). This means that little additional entropy is gained by using both quadrature components if the wavelets selected are narrowband.

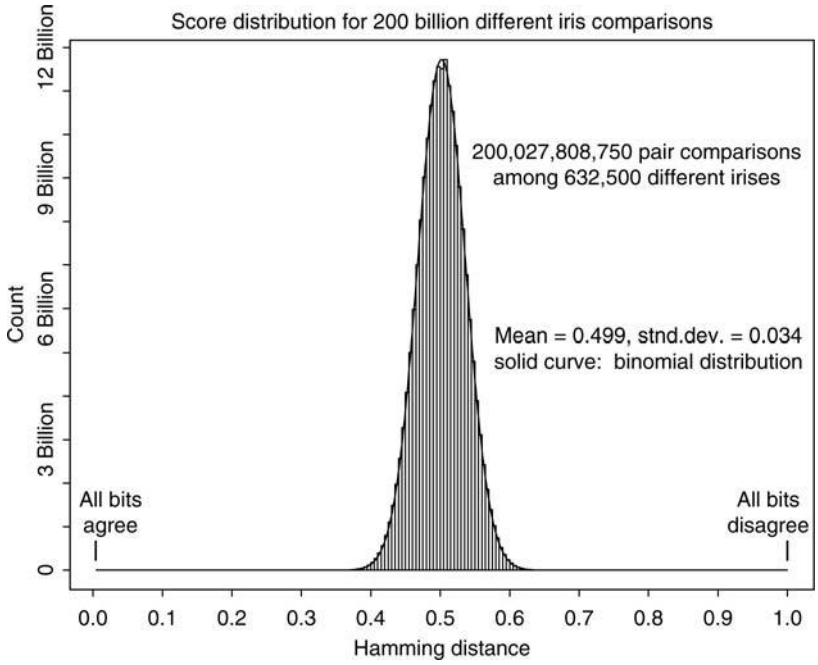
other corruptions), and one should expect the real and imaginary bit pairs to be independent because of the orthogonality of the corresponding parts of the Gabor wavelets. Therefore, as with any sequence of "tosses" from two independent and fair "coins," one should expect a 50%-50% level of agreement between the bits just by chance. For quadrature IrisCode bit pairs, this is confirmed in the upper panel of Fig. 5, showing the frequency with which different ▶ Hamming Distances (proportion of disagreeing bits) were observed between the corresponding real and imaginary parts of 545 IrisCodes, computed over all pairs of {Re, Im} corresponding bits. With a mean Hamming Distance of 0.501 ± 0.011, the expected finding of independence between such equiprobable bits is clearly observed. However, because the wavelets used to compute these IrisCodes had relatively narrow bandwidths, a strong degree of *phase coherence* is present in their outputs. The consequence of such phase coherence is that the real and imaginary parts become highly correlated under a shift.

The lower panel of Fig. 5 plots a histogram of Hamming Distances observed between the real and imaginary bit streams after one stream has been shifted by $\pi/2$ relative to the other. Now we see that these bit streams are far from independent. Instead, with an average probability of 0.912, they are simply complements of each other. The cause of this effect is clear from the fact that narrowband wavelets (encompassing many cycles) are almost equivalent to each other, or negatively so, when shifted by $\pm \pi/2$. In the case illustrated by Fig. 5, the negative correlation is so strong that there is little justification for using both sets of bits if they are computed using relatively narrowband wavelets; almost no additional entropy (or information) is gained. Of course, this does not apply to wavelets having broader bandwidth, when, as noted in Fig. 3, the relationship is more like that between the first and second derivatives. In that case, if one were forced to choose one over the other, the second derivative (corresponding to the real part of a Gabor wavelet) would be the better choice, because the first derivative is sensitive to gradients of illumination, as may often occur in iris recognition systems using off-axis illumination.

In the version of this algorithm that is currently used in all public deployments of iris recognition worldwide, the wavelet parameters were chosen to optimize operation in *identification mode,* which requires exhaustive search through enrolled databases without succumbing to False Matches despite the large numbers of possibilities. The benefit of operating in this mode is that users need not assert their identities, as would be required by operation in *verification mode* in which only a one-to-one comparison is done against a single identity asserted by, for example, a token or card. But successful operation in identification mode requires that the distribution of similarity scores obtained when different irises are compared must be confined by rapidly attenuating tails, since that distribution is effectively being sampled a large number N times when searching a database where the number N of stored IrisCodes might correspond to the size of a nation's population. The larger the number of samples N, the greater the likelihood of finding a sample far out along the tail and thus a possible False Match. Figure 6 shows the result of 200 billion iris cross-comparisons obtained from one such national border-crossing deployment at all air, land, and seaports of entry into the United Arab Emirates. Since comparisons between different persons never generate Hamming Distance (dissimilarity fraction) scores smaller than about 0.25, at least among these 200 billion such comparisons, we see that successful recognition using this biometric requires only that different images of a given iris are of sufficient quality that no more than about 25% of their computed IrisCode bits disagree. Under reasonable image acquisition conditions, this is easily achieved.

Gabor Wavelets in Other Biometrics

A powerful advantage of the Gabor wavelet approach to iris encoding and recognition is its great speed. The complete execution time for all aspects of the image processing, starting with a raw image, including the localisation of the iris, detection of all boundaries including eyelids and their exclusion, detection and removal of eyelashes and other noise, normalization in a dimensionless coordinate system, and demodulation and compilation of the IrisCode with its masking bits, is less than 30 ms on a 3 GHz processor. This speed means that more than 30 complete image frames can be fully processed per second, and so the process can operate at the same rate as the video frame rate itself. Of the 30 ms consumed per image frame, the vast majority of processing time is spent on localization, segmentation, and normalization operations; less than 1 ms is consumed by demodulation with the Gabor



Iris Encoding and Recognition using Gabor Wavelets. Figure 6 Distribution of Hamming Distance scores (fraction of disagreeing bits) obtained in 200 billion cross-comparisons among 632,500 different iris patterns enrolled in the United Arab Emirates border-crossing deployment. The rapid attenuation of the left tail means that False Matches are avoided even in exhaustive searches through national databases, provided that the decision policy allows no more than about 25% of the bits to disagree (HD $<$ 0.25) when declaring a match.

wavelets and creation of the IrisCode. Once an IrisCode has been computed, the simplicity of the comparison process and decision algorithm allows databases to be searched at the speed of about 1 million IrisCodes/second per 3 GHz processor.

Since these execution speeds for image processing and for matching are very favorable compared to those of other biometrics, efforts have been made to adapt these methods for the other modalities as well. Notable among these are face [13], fingerprint [14], and palmprint [15] recognition. Besides the speed advantage, and the design benefits of formulating a biometric recognition task as a test of statistical independence on the outputs of logico-linear operators, the mathematical merits of Gabor wavelets as reviewed in this chapter also contribute fundamentally to the success of this framework for image coding and pattern recognition.

Related Entries

► Active Contours in Iris Recognition
► Iris-on-the-Move™

References

1. Daugman, J.G.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. **15**, 1148–1161 (1993)
2. Daugman, J.G.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**, 21–30 (2004)
3. Meyer, Y.: Principe d'incertitude, bases hilbertiennes et algebres d'operateurs. Bourbaki Seminar 662 (1985)
4. Daubechies, Y.: Orthonormal bases of compactly supported wavelets. Comm. Pure Appl. Math. **41**(7), 909–996 (1988)
5. Gabor, D.: Theory of communication. J. Inst. Electr. Eng. **93**, 429–457 (1946)
6. Hubel, D.G., Wiesel, T.N.: Sequence regularity and geometry of orientation columns in the monkey striate cortex. J. Comp. Neurol. **158**, 267–293 (1974)
7. Jones, J.P., Palmer, L.A.: An evaluation of the 2D Gabor filter model of simple receptive fields in cat striate cortex. J. Neurophysiol. **58**, 1233–1258 (1987)
8. Pollen, D.A., Ronner, S.F.: Phase relationships between adjacent simple cells in the visual cortex. Science **212**, 1409–1411 (1981)
9. Daugman, J.G.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimised by two-dimensional visual cortical filters. J. Opt. Soc. Am. A **2**, 1160–1169 (1985)

10. Daugman, J.G.: Complete discrete 2D Gabor transforms by neural networks for image analysis and compression. IEEE Trans. Acoust. Speech Signal Process. **36**, 1169–1179 (1988)
11. Newton, I.: Method of fluxions. Manuscript in Trinity College Library, University of Cambridge (1671)
12. Oppenheim, A.V., Lim, J.S.: The importance of phase in signals. Proc. IEEE **69**, 529–541 (1981)
13. Wiskott, L., Fellous, J.M., Kuiger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**, 775–779 (1997)
14. Jain, A.K., Prabhakar, S., Hong, L., Pankanti, S.: Filterbank-based fingerprint matching. IEEE Trans. Image Process. **9**, 846–859 (2000)
15. Kong, A.W.K.: Palmprint Iientification based on generalization of irisCode. Ph.D. thesis, University of Waterloo, ON, Canada (2007)

Iris Image Capture Device

▶ Iris Acquisition Device
▶ Iris Device

Iris Image Data Interchange Formats, Standardization

James L. Cambier
Crossmatch Technologies, RCA Blvd, FL, USA

Synonyms

Iris interchange format standards; Iris data interchange standards

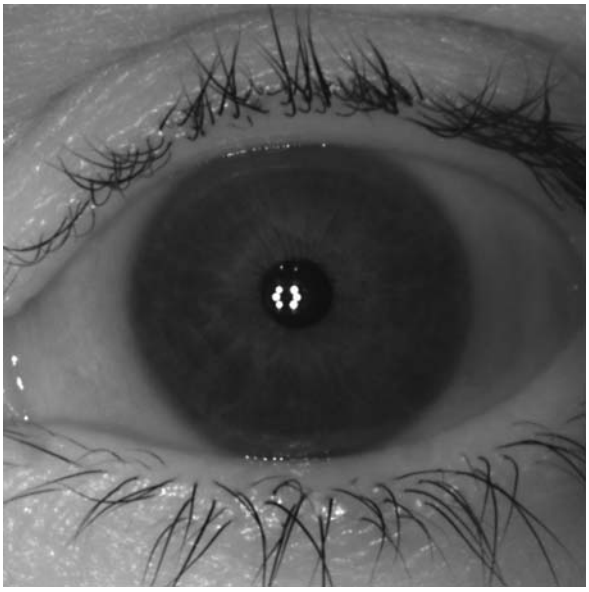
Definition

Iris recognition is a biometric technology that uses the unique, stable, and repeatable texture patterns observed within the iris of the human eye, the colored annular ring that surrounds the pupil. Iris recognition systems typically consist of specialized cameras and software that processes images of the eye to extract and encode iris features in a template, and match the presented iris templates to those in a database to identify the individual. Applications include controlled access to buildings, border security, trusted traveler programs, and authentication of emergency aid, entitlement, and citizen benefit recipients. Iris image interchange standards have been developed to facilitate the exchange of iris image data among diverse cameras, processing algorithms, and biometric databases. Existing standards include ANSI INCITS 379 Iris Image Interchange Format and ISO/IEC 19794-6 Information technology: Biometric data interchange formats – Part 6: Iris image data.

Introduction

The human ▶ iris is a colored annular ring that surrounds the ▶ pupil, a variable aperture that admits light to form an image on the ▶ retina, the light-sensitive surface in the back of the eye. The iris is a muscular structure that contains a variety of texture features, including pits, furrows, and radial striations (Fig. 1). The rich and unique texture of the iris has long been recognized, and a number of biometric systems based on the iris have been described [1–4]. Particular algorithmic approaches to the extraction, encoding, and matching of iris texture features have also been published [5–8].

With the advent of multiple vendors of iris cameras and systems, and various algorithmic approaches to



Iris Image Data Interchange Formats, Standardization.
Figure 1 Iris image.

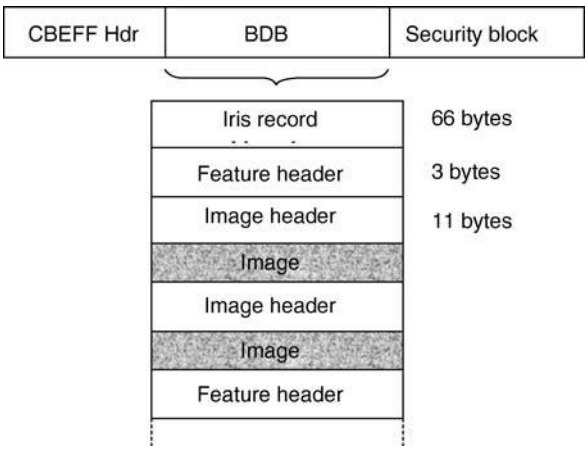
iris-based recognition, it became clear that the development of data interchange standards for iris images would (1) facilitate the exchange of iris images among multiple vendors, applications, and algorithms, (2) enable the compilation of iris image databases for comparative testing and evaluation of multiple algorithms, and (3) support the evolutionary development of new iris algorithms by preserving existing enrollment databases in the form of reusable images.

The development of iris standards has been a collaborative effort of a number of biometric vendors, government agencies, and academic institutions, and has resulted in both a US standard, ANSI INCITS 379 Iris Image Interchange Format [9], and an international standard, ISO/IEC 19794-6 Information technology: Biometric data interchange formats – Part 6: Iris image data [10]. The ANSI standard was developed first and became the basis for the international standard; as a result the two are virtually identical.

The iris standards support two different data formats, a ▶ rectilinear format in which the iris image is represented in standard Cartesian $(x - y)$ coordinates, and a ▶ polar format, in which the (approximately) circular iris is represented in polar $(r - \theta)$ coordinates. The rectilinear format provides the highest interoperability, while the polar format retains only the area of specific interest, the iris, and thus provides a more compact representation. The standard allows a number of different image intensity representations (color, monochrome, etc.), compression schemes, and geometric orientations. Finally, an appendix to the standard contains a set of recommendations for iris image capture, addressing quality metrics, resolution, illumination, distortion, noise, orientation, and other properties.

Data Formats

The iris image record is a nested structure that contains several headers and one or more images (Fig. 2). The overall structure consists of a CBEFF header [11], the Biometric Data Block (BDB), and a Security Block (SB). The CBEFF header contains information about image quality, the origin of the BDB format used, and information about the type of biometric contained in the data record. The BDB contains an Iris Record Header, one or two Feature Headers, and one or more images, each preceded by an Image Header. The Record Header contains information and parameters specifying the format of all of the images in the record, such as geometric format (rectilinear or polar), orientation, dimensions, number of intensity levels, number of intensity bands (i.e., monochrome, color, etc.), and compression methods. The Feature Header indicates which eye was imaged (left or right), if known, and the number of images recorded for that eye. Finally the Image Header contains an image sequence number, image quality value, size of the image data, and information about the rotational position of the iris, if known.



Iris Image Data Interchange Formats, Standardization.
Figure 2 Iris image data record.

The rectilinear image format is a conventional image composed of rows and columns where each entry corresponds to one pixel (picture element) produced by the image sensor. If the image contains multiple color bands, such as red-green-blue, each pixel is recorded as three sequential values.

An image in polar format is produced by processing the rectilinear image to find the iris center, pupil boundary, and outer iris boundary or ▶ limbus. The portion of the image containing the iris is then sampled along radial lines emanating from the iris center and extending from the pupil boundary to the iris boundary at particular angles. The result is an image in polar coordinates. It is stored as a matrix in which each column corresponds to one angular orientation θ and each row corresponds to a particular radial distance r. Interoperability of polar images may be limited by the accuracy and consistency with which

the pupil and iris boundaries are determined. Their advantage is very compact iris data representation, since no image information within the pupil or outside the iris is included.

Image Properties

The properties recorded for rectilinear and polar images differ to some extent. The properties are as follows, with these differences noted:

Image orientation – the images may be recorded in "canonical" form, in which the top of the eye is at the top (first row) of the image and, for a right eye, the nasal side of the eye (that closest to the nose) is on the right side of the image. Alternatively, the image may be flipped vertically or horizontally. For polar images the orientation refers to the rectilinear image used to produce the polar image.

Scan type – the images may have been collected using progressive scanning, in which each row is captured in sequence, or interlaced scanning, in which all odd rows are captured followed by all even rows. Note that in the latter case the image is still stored in strict row sequence.

Data format – the images may be uncompressed (or "raw") or compressed; they may be color or monochrome, and if compressed the applicable compression standard is referenced.

Image size – the image dimensions are recorded as width and height, recognizing that for polar format the width corresponds to angular samples and the height to radial samples. The number of bits allocated to each intensity value is also recorded.

Rotation angle – relative rotation between enrollment and recognition images must be either corrected or accommodated in the match process by searching over a range of rotations [5, 7] or using templates based on rotation-invariant features [3]. Some cameras are capable of approximating the ▶ rotation angle by capturing both eyes simultaneously and calculating the angle of the interpupillary line with a horizontal reference. This angle information may be recorded.

Camera information – space is allocated in the image properties to record a unique identifier for the camera and the date and time of capture.

Occlusion marking – local areas of the iris may be occluded by reflections, eyelids, or eyelashes and therefore, should not be used to generate template information. The standard includes fields for recording whether occlusions have been detected, and if so how they are marked in the image data (usually as a reserved intensity value).

Image Quality

The iris standard allows the originator of an iris image to indicate the quality of the image on a scale from 1 to 100. The interpretation of the quality score is at the discretion of the originator, but the following general quality interpretations are recommended:

1–25 – unacceptable quality
26–50 – low quality, suitable for verification in low-cost systems
51–75 – medium quality, suitable for verification identification (one to many matching) in medium security applications
76–100 – highest quality images, suitable for enrollment

Image Capture Recommendations

Appendix A of the standard provides specific recommendations on capture of iris images, based on vendor experience in commercial deployments of iris recognition. These recommendations include the following:

Resolution
Grayscale range
Illumination wavelength
Contrast
Iris visibility
Pixel aspect ratio
Image scale
Optical distortion
Noise content
Image orientation
Subject presentation

ANSI and ISO Differences

Although the American National standard and the international standard are more or less identical there

Iris Image Data Interchange Formats, Standardization. Table 1 Comparison of ANSI and ISO versions of iris image interchange standards

Attribute	ANSI INCITS 379	ISO/IEC 19794-6
Image quality field description	Described as 4 categories with numerical range 1–100	Described as value with numerical range 1–100
Second level header title	Feature Header	Biometric Subtype Header
CBEFF Product Identifier	Contained in Iris Record Header	Contained in CBEFF Header
User Identification No.	Contained in Iris Record Header	Contained in CBEFF Header

are also a number of minor differences. These are summarized in Table 1.

Iris Standards Adoption

A number of national and international governments and organizations have officially adopted iris image data standards for current and planned programs. Within the US Department of Homeland Security the Transportation Security Agency (TSA) has adopted ISO/IEC 19794-6, in addition to ANSI INCITS standards for Finger Minutiae, Face Recognition, and the Common Biometric Exchange Formats Framework (CBEFF) for its Registered Traveler program. In this program the iris rectilinear format is used for transmission of enrollment images to a central data center, and polar format is used for storage of iris image data on the registered traveler card. The International Civil Aviation Organization (ICAO) has adopted ISO/IEC standards for face, finger, and iris biometrics, in addition to the ISO/IEC CBEFF standard, for its Machine Readable Travel Documents (MRTDs).

Current Iris Standards Activity

Current standards activities related to the iris image standards include development of conformance test standards, development of revisions of the current standards, and adoption of the international ISO/IEC standard by various national standards organizations.

Conformance testing is intended to assess commercial products that claim to support the iris standard. Developers of software applications and biometric products may interpret the standard differently from one another, and as a result their implementations of the specification may differ and not interoperate. Conformance to the standard is a necessary prerequisite for achieving interoperability among implementations; therefore, there is a need for a standardized, generally accepted, conformance testing methodology that would allow implementation of a set of test tools realizing this methodology. The conformance test standards are applicable to the development and use of conformity test method specifications, conformity test suites for Iris Image Data Record requirements as specified the ANSI and ISO standards, and conformance testing programs for conforming products. They are intended primarily for use by testing organizations, but may be applied by developers and users of test method specifications and test method implementations.

The international biometric standards committee ISO/IEC JTC1 SC37 is currently developing a revised version of ISO/IEC 19794-6 to address needed clarification of the original standard, incorporate certain technology innovations, and respond to new customer requirements. In particular, SC37 is considering both removal of the requirement for mandatory embedding of the iris image record within a CBEFF data structure and elimination of the polar image format.

International standards such as ISO/IEC 19794-6 may be adopted by national standards bodies for use as national standards. The U.S. has adopted the international iris standard as a US standard, and has withdrawn ANSI INCITS 379, cancelling the revision project for this standard and the conformance testing methodology standard project.

Summary

The published ANSI INCITS and ISO/IEC standards for iris image data interchange were developed in a cooperative, collaborative process with participants from numerous commercial entities, government agencies, and

academic institutions from the US and other countries. The standards have been successfully adopted by a number of US and international organizations that desire to use iris recognition in operational deployments. The standards are sufficiently flexible to accommodate the products of diverse commercial vendors and user organizations. Ongoing standards development work will result in more flexible and interoperable standards and conformance test suites that may be used to effectively test the compatibility of various products with the standards.

Related Entries

► Biometric Sample Quality, Standardization
► Biometric Technical Interface, Standardization
► Common Biometric Exchange Framework Formats, Standardization
► Conformance Testing Methodologies for Biometric Data Interchange Formats, Standardization of
► Data Interchange Format, Standardization
► Iris Image Quality
► Iris Recognition, Overview

References

1. Flom, L., Safir, A.: Iris recognition system, US Patent No. 4,641,349, United States Patent and Trademark Office, Washington DC, (1984)
2. Kim, D., Ryoo, J.: Iris identification system and method of identifying a person through iris recognition, US Patent No. 6,247,813, United States Patent and Trademark Office, (2001)
3. Monro, D., Rakshit, S., Zhang.: DCT-Based Iris Recognition. IEEE Trans. Pattern Anal. Mach. Intell. **29**(4), 586–595 (2007)
4. Wildes, R., Asmuth, J., Hanna, K., Hsu, S., Kolczynski, R., Matey, J., McBride, S.: Automated, non-invasive iris recognition system and method, US Patent No. 5,572,596, United States Patent and Trademark Office, (1996)
5. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. **15**(11), 1148–1160, (1993)
6. Ma, L., Wang, Y., Tan, T., Zhang, D.: Personal identification based on iris texture analysis. IEEE Trans. Pattern Anal. Mach. Intell. **25**(12), 1519–1533, (2003)
7. Masek, L.: Recognition of human iris patterns for biometric identification, Bachelor of Engineering thesis, School of Computer Science and Software Engineering, The University of Western Australia, (2003)
8. Tisse, C., Martin, L., Torres, L., Robert, M.: Person identification technique using iris recognition, In: Proceedings of the 15th International Conference on Vision Interface, Calgary, Italy, pp. 294–299 (May 29, 2006)
9. ANSI INCITS 379-2004 Iris image interchange format. American National Standards Institute, (2004)
10. ISO/IEC 19794-6: 2005 Information technology – Biometric data interchange formats – Part 6: Iris image data, International Standards Organization, (2005)
11. ISO/IEC 19785-1:2006 Information technology – Common biometric exchange formats framework – Part 1: Data element specification, International Standards Organization, (2006)

Iris Image Enhancement by Super-Resolution Method

► Iris Super-Resolution

Iris Image Quality

Natalia A. Schmid
West Virginia University, Morgantown, USA

Synonyms

Information content of iris images; Iris quality metrics

Definition

Iris image quality evaluation is a procedure of measuring information content of iris imagery at the stage of iris acquisition or at early processing stage. The information content may be decided to be insufficient to be used for iris identification based on a single image. In this case, the image may be discarded, or combined with other imagery to improve recognition capabilities of an iris system. Evaluated quality metrics would be the guidelines in making decisions regarding further steps with respect to acquired imagery.

Introduction

Iris image quality assessment is an important research thrust recently identified in the field of iris biometrics [1–3]. This research is tightly related to the research on

▶ nonideal iris. Its major role is to determine, at the stage of data acquisition or at the early stage of processing, what the amount of information for the purposes of processing, recognition, and fusion this imagery contains. Is it informative enough for performing further processing steps or should be discarded? Is it informative enough for being combined with other images and result in improved recognition performance? The quality metrics play an important role in automated biometric systems for three reasons: (1) system performance (segmentation and recognition), (2) interoperability, and (3) data enhancement.

The quality metrics play an important role in automated biometric systems for two reasons: (1) system performance (segmentation and recognition), and (2) interoperability.

A traditional approach in evaluating iris image quality is to identify a single or a sequence of physical phenomena that influences formation of query iris imagery at the image acquisition stage (see the entry on Image Acquisition). The distortions that identified physical phenomena introduced are then modeled mathematically. To evaluate the level of distortion present in an iris image, a single or a set of metrics or quality measures is specified. The metrics can be absolute measures or relative measures. The absolute metrics do not assume comparison of query image with a reference image. The relative metric measures the presence of some distortions with respect to a specified reference image.

The following sections provide a short survey of the literature on iris image quality, introduce a set of iris quality measures and suggest a number of techniques to combine individual quality measures in a single score.

Survey of Iris Quality Metrics

Previous work on iris image quality can be placed in two categories: local and global analysis. Zhu et al., [4] evaluate quality by analyzing the coefficients of particular areas of iris's texture by employing discrete wavelet decomposition. Chen et al., [5] classify iris quality by measuring the energy of concentric iris bands obtained from 2-D wavelets. Ma et al., [6] analyze the Fourier spectra of local iris regions to characterize out-of-focus and motion blur and occlusions. Zhang and Salganicaff [7] examine the sharpness of the region between the pupil and the iris. Daugman [8] and Kang and Park [9] characterize quality by quantifying the energy of high spatial frequencies over the entire image region. Belcher and Du [10] propose a clarity measure by comparing the sharpness loss within various iris image regions against the blurred version of the same regions. The major feature of these approaches is that the evaluation of iris image quality is reduced to the estimation of a single [5, 7, 8, 9] or a pair of factors [6], such as out-of-focus blur, motion blur, and occlusion.

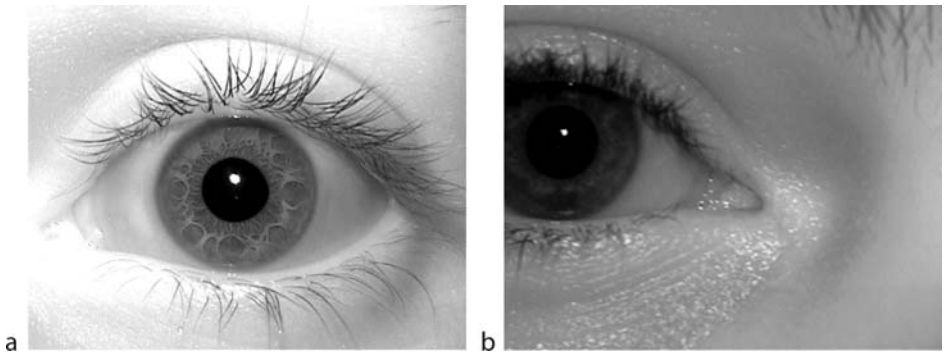
A broader range of physical phenomena that can be observed in nonideal iris imagery was characterized by Kalka et al., [11, 12]. The proposed factors include out-of-focus and motion blur, occlusion, specular reflection, illumination, off-angle, and pixel count. The strength of the phenomena and its influence was evaluated through modified or newly designed iris quality metrics. These factors based on the extensive analysis carried out by the authors affect the segmentation and ultimately recognition performance of iris recognition systems. An example of two iris images from ICE dataset [13] and their corresponding pentagram plots are displayed in Figure 1 and Figure 2. In a pentagram, each axis represents a quality metric. The quality score is normalized to take values between zero and one. The value one corresponds to the lowest quality, the value zero is the highest quality.

Since most of quality metrics contain some common information (for example, motion and out-of-focus blurs are physically related; occlusion and pixel counts usually contain redundant information; illumination and contrast are also related), they have to be estimated jointly. However, all recently designed quality assessing algorithms treat individual quality metrics as independent and thus evaluate them separately.

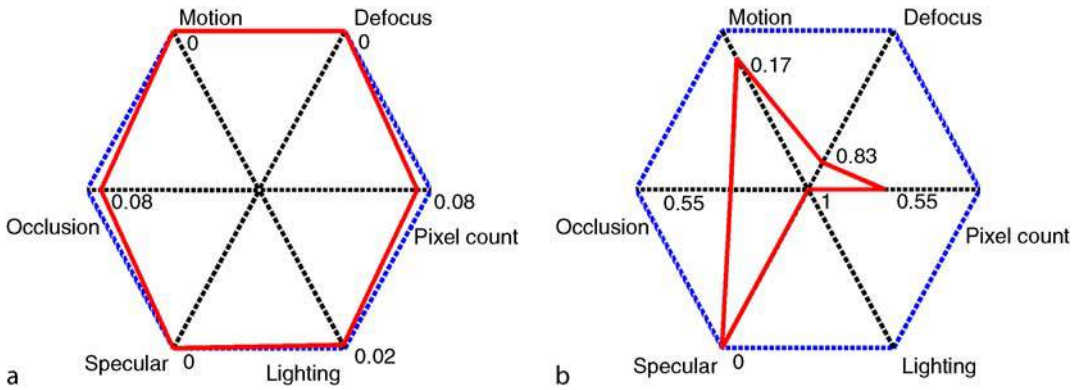
While any processing reduces information content of quality metrics, in some applications it is beneficial to have a single quality score that characterizes the overall image quality. In this case we would appeal to rules of combining scores.

Rules of Combination

The quality factors (metrics) can be used individually or combined into a single score through a simple static or an adaptive rule. Among static rules the simple sum rule is a computationally efficient method. More complex (adaptive) rules such as Bayesian,



Iris Image Quality. Figure 1 Sample images from ICE dataset. **(a)** Good quality image. **(b)** Poor quality image.



Iris Image Quality. Figure 2 Pentagram plot of quality for the images in Figure 1.

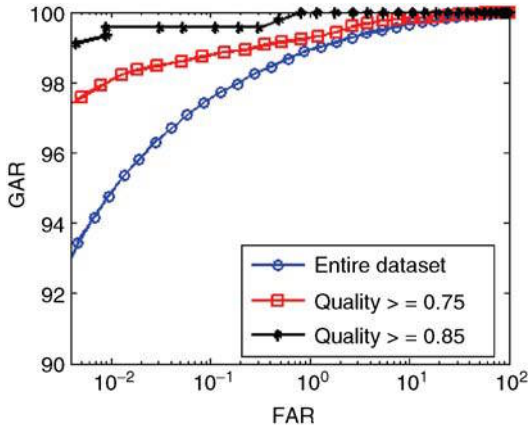
Dempster-Shafer, weighted Sum, or any previously designed fusion strategy to combine classifiers can also be used to combine quality metrics into a single score. These rules are more fundamental and flexible, but require intensive computations.

Relevance to Recognition Performance

It is well understood that the capabilities of various designed metrics to evaluate the quality of iris images has to be related to their capabilities to predict recognition performance by quickly analyzing the quality of imagery. Since in most analyzed cases recognition performance is nonlinearly related to quality metrics, it is hard to evaluate precision with which quality metrics and recognition performance are related. It is because of this reason most of the designed quality metrics are not highly correlated with one another.

Figure 3 demonstrates the relationship between the iris quality metrics evaluated by following the procedure in Kalka et al., and combined through application of Dempster-Shafer rule and the recognition performance. The results are obtained using images from ICE dataset [13]. To obtain the plots in Figure 3, each image undergoes quality evaluation. The images are ranked based on their quality score. The receiver operating curves are plotted for three cases. In the first case, the entire dataset is used. In the second case, the images with the quality above the score 0.75 are selected. In the last case, the images with the quality above the score 0.85 are selected. Note that here the score value one corresponds to the highest quality. From Figure 3, the combined metric predicts recognition performance relatively well.

This indicates that the information extracted fromiris imagery is correlated to some degree with the recognition performance. The question that remains to be addresses if it is possible to establish an explicit



Iris Image Quality. **Figure 3** Establishing the relationship between estimated quality metrics and verification performance of an iris recognition system. The results are based on applications of a Gabor-filter based encoding algorithm to iris images from ICE dataset.

relationship between quality metrics and recognition (verification) performance.

In practice most of metrics are designed to measure signal-to-noise ratio, ratio of powers, directional power, sharpness of edges, ratio of pixels, empirical entropy and relative entropy. These metrics are not explicitly related to probability of recognition error or receiver operating characteristic (ROC) curve, two traditional measures of recognition and verification performance. Furthermore, the degree of nonlinearity and nonlinear model relating the quality metrics and recognition performance are hard to evaluate.

The conclusions above are supported by the results of an extensive research published in the literature on the topic of feature selection. The problem of feature selection can be briefly summarized as follows. Due to suboptimality of an image/signal encoding procedure (feature extraction procedure) templates contain a certain number of noisy components or components uninformative for pattern recognition. Removal of these components, therefore, often leads to improved recognition performance. In the past, some substantial efforts have been made to find an information measure that would be capable of evaluating the information content of a feature and, at the same time, would be able to predict changes in recognition performance due to removal of a feature [14, 15]. In spite of long term efforts, the final conclusions may not look worth of these efforts: (1) there is a single family of statistical models (Gaussian family) that allows establishing an explicit (exponential) relationship between the information content of a feature and the recognition performance. Unfortunately for biometrics, only few types of biometric templates (encoded data) can be modeled as being Gaussian distributed. (2) The traditional measures of information such as signal-to-noise-ratio, entropy, relative entropy, and mutual information, are only asymptotically related to probability of recognition error [16]. When these measures are adapted to the problem of empirical evaluation of the information content of a feature, the relationship between the empirically evaluated measures and the probability of recognition error can be established only under the condition that an increasing amount of data can be used to evaluate information in features and provided that the empirical information measures converge in some sense to the true measures.

Summary

Iris images quality metrics provide us with a fast way to predict information content of biometric data. Current quality metrics are designed to evaluate quality factors individually, that is, separately. These quality factors may further be combined by invoking static or adaptive rules applied to individual scores. Since various quality factors contain redundant information about one another, a more optimal solution to the problem of evaluation of iris image quality would be to evaluate the quality factors jointly. Also, identifying the most important quality factors that influence the recognition performance remains an open problem.

Related Entries

- ▶ Iris Segmentation
- ▶ Iris Recognition
- ▶ Score Fusion and Normalization Rules

References

1. Biometric quality workshop. national institute of standards and technology. (2006)
2. Biometric quality workshop ii. national institute of standards and technology. (2007)
3. Bowyer, K., Hollingsworth, K. Flynn, P.: Image understanding for iris biometrics: A survey. J Comp Vision and Image understanding. **110**, 281–307 (2007)

4. Zhu, X., Liu, Y., Ming, X., Cui, Q.: A quality evaluation method of iris images sequence based on wavelet coefficients in region of interest. In: Proceedings of the fourth International Conference on Computer and Information Technology, pp. 24–27 (2004)
5. Chen, Y., Dass, S., J.A.: Localized iris quality using 2-d wavelets. In: Proceedings of International Conference on Biometrics, pp. 373–381. Baltimore, MD (2006)
6. Ma, L., Tan, T., Zhang, Y.W.D.: Personal identification based on iris texture analysis. IEEE Trans. Pattern Anal. Mach. Intell. **25**(12), 1519–1533 (2003)
7. Zhang, G., Salganicoff, M.: Method of measuring the focus of close-up image of eyes (1999)
8. Daugman, J.: How iris recognition works. IEEE Trans. Circ. Syst Video Technol. **14**(1), 21–30 (2004)
9. Kang, B., Park, K.R.: A study on iris image restoration. In: Proceedings Audio Video Based Personal Authent., vol. 3546 (2005)
10. Belcher, C., Du, Y.: Information distance based selective feature clarity measure for iris recognition. In: Proceedings SPIE Symposium on Defense and Security. Conference on Human Identification Technology IV., vol. 6494 (2007)
11. Kalka, N.D.: Image quality assessment for iris biometric. https://eidr.wvu.edu/files/4447/Kalka_Nathan_thesis.pdf (2005)
12. Kalka, N., Zuo, J., Dorairaj, V., Schmid, N., Cukic, B.: Image quality assessment for iris biometric. In: Proceedings of 2006 SPIE Conference on Biometric Technology for Human Identification III. Orlando, FL (2006)
13. Liu, X., Bowyer, K.W., Flynn, P.J.: Iris recognition and verification experiments with improved segmentation method. In: proceedings Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID) (2005)
14. Ben-Bassat, M.: Use of distance measures, information measures and error bounds in feature evaluation. Handbook of Statistical, P.R. Krishnaiah and L.N. Kanal, North-Holland Publishing Company, pp. 773–791 (1982)
15. Jain, A.K., R.P.W.D., Muo, J.: Statistical pattern recognition: A review. IEEE Trans on Pattern Analysis and Machine Intelligence. **22**, 4–36 (2000)
16. Cover, T.H., Thomas, J.A.: Elements of Information Theory. Wileg Series in Telecommunications, New York (1991)

Iris Interchange Format Standards

► Iris Image Data Interchange Formats, Standardization

Iris Localization

Iris localization is one important stage in iris recognition system. The goal of iris localization is to find the boundary between iris and sclera and between iris and pupil. Because the shape of pupil and the shape of iris look very close to a circle, most of the iris localization algorithm tries to perform circle fitting on the eye image. In this way, the results of iris localization can be expressed as two circles.

► Automatic Classification of Left/Right Iris Image
► Iris Super-Resolution

Iris on the Move™

James R. Matey
Electrical and Computer Engineering Department, US Naval Academy, Annapolis, MD, USA

I

Synonyms

Drive-up; Iris at a glance; Minimal constraint iris recognition; Portal; Walk-through; Walk-up

Definition

Iris on the Move™, commonly referred to as IOM, is an approach to acquiring iris images suitable for iris recognition while minimizing the constraints that need to be placed on the subject. Sarnoff Corporation developed IOM under contract to the United States Government: NMA401-02-9-2001.

Multiple implementations of IOM have been demonstrated; they include: a portal walk-through system, an on-the-wall walk-up system, an over-the-door walk-through system and a roadside drive-up system. The key features of the IOM systems are large ► capture volume, large ► standoff distances and the capability of capturing recognition quality iris images while the subject is in motion [1].

The Panasonic "Iris at a Glance" systems and the H-box™ from the Hoyos Group are examples of related technologies.

Introduction

Iris recognition is one of the strongest biometric available [2–4]. Iris recognition is a strong biometric

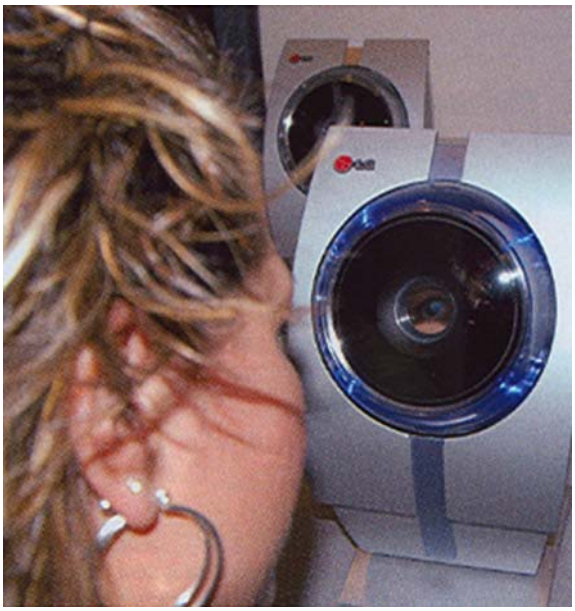
because: (1) the human iris is a complex structure with a high degree of randomness; (2) the iris is protected; (3) the iris is accessible; and (4) the structures of the iris that are used for iris recognition are stable, from early childhood – in the absence of illness or injury that disrupts the iris tissue.

The first mention of iris patterns as a biometric was likely a paper by Bertillon [5]; several others subsequently suggested iris patterns as a biometric and the idea was a plot element in the 1983 James Bond film *Never Say Never Again* [6]. However, it was not until the early 1990s that John Daugman developed a practical algorithm for iris recognition based on Gabor wavelets [7]. Minor variants on the Daugman algorithm remain the dominant algorithms in commercial iris recognition systems as of 2007, though there are vigorous research efforts into alternative algorithms [8]. The commonly used name for the Daugman algorithms in current use is iris2pi.

The capability of providing extremely low ► false match rates is one of the great strengths of iris2pi. Daugman has published credible evidence that iris2pi can support false match rates of the order of a part in a trillion [9]. However, the quality of the iris images required for such results is high. The standards for iris images published by the ISO [10] call for 100–200 pixels across the iris and 40 dB signal-to-noise-ratio (SNR). In the decade following the development of the Daugman algorithm, Iridian, IrisGuard, LG, OKI, Panasonic, and Securimetrics all brought iris recognition cameras to market that could provide such high quality iris images. Figure 1 is an example of one of these cameras that provides the full 200 pixels across the iris.

All these systems require substantial cooperation from the subjects and impose significant constraints on them. In general, imposing constraints and requiring cooperation reduces the ease of use of any biometric system. This may not be a significant issue for ► habituated subjects, but it is a concern for nonhabituated subjects – subjects who do not use the system on a routine basis.

Iris on the Move™ was developed to relax subject constraints and make it possible for subjects to use an iris recognition system with little instruction. The constraints of greatest interest are capture volume, standoff, ► residence time, subject motion, effects of ambient illumination, and overall ease of use. Capture volume is the three dimensional volume throughout which the iris camera is capable of acquiring an acceptable image. Standoff is the distance from the camera to the subject; in some systems there may be two standoff distances – subject to camera and subject to illumination. Residence time is the length of time that a subject must stay within the capture volume to enable collection of an acceptable image. Subject motion within the capture volume can be characterized in terms of a velocity – the maximum velocity at which a subject can move during collection of an acceptable image. Overall ease of use can be characterized in terms of the complexity of the instructions that must be provided to an uninitiated subject to insure acceptable iris image collection on their first use of a system.



Iris on the Move™. Figure 1 An enrollment quality iris image capture device, the LG IrisAccess 3000. (Photo provided courtesy of Mohammed Murad of LG Electronics USA Inc., Iris Technology Division.)

The reader may well ask, "What is an *acceptable* iris image?". Operationally an acceptable image is a recognition image that matches well against an enrollment image taken under ideal circumstances – the enrollment image meets ISO standards, highly constrained subject, no time limit, repeated trials available to obtain optimum quality image, and close supervision of the enrollment process by a well trained operator. It is important to note that iris2pi looks at the local phase of the iris image in predefined spatial frequency bands (see the entry on Gabor wavelet iris encoding). Hence, the

image quality metrics that are used for assessing the visual quality of images in daily life do not map directly to the quality metrics for an image used for iris recognition with iris2pi. For example, a reduction in the contrast of a properly adjusted television display will almost always reduce its visual quality – while a reduction in the contrast of a good quality iris image may have little impact on its match score against another good image of the same iris. This can be understood in mathematical terms: phase and amplitude are orthogonal/independent coordinates. Human vision is more sensitive to amplitude than phase; iris2pi is more sensitive to phase than amplitude.

IOM Design Considerations

The single most important IOM design consideration is the exploitation of asymmetries between enrollment and recognition. As Daugman points out [9], the number of degrees of freedom (as measured by the number of valid bits in an iris2pi template) in an iris template can vary depending on the quality of underlying image. Fewer degrees of freedom translate to a broader ▶ imposter distribution. This broadening is taken into account in many systems by adjusting the Hamming distance of a match based on the number valid bits compared between two templates to maintain a constant false match rate for a fixed match criterion in the presence of such variation. At enrollment, significant effort can be expended, once per subject, to capture a pristine image of the iris. Pristine enrollment images enable us to purge the database of duplicates – even when the database is large. They also ensure that maximal use is made of the degrees of freedom available in a lower quality recognition image.

If there is high quality enrollment, at recognition, it can afford a reduction in the quality of the image. The ▶ authentics distribution will be dominated by the quality of the recognition images. Reduction in the quality of the recognition images *will* broaden the authentics distribution and increase the ▶ false non-match rate (FNMR). However, it is possible to adjust for the increase in the single attempt FNMR through use of multiple attempts.

IOM systems use high quality images from iris cameras equivalent to that in Fig. 1 to build a strong iris database. At each recognition attempt IOM systems take multiple, lower quality images of the subject iris and compare them against the database. Let FMR(1) be the false match rate for a recognition attempt where only one image is collected, and FNMR(1) corresponding false non-match rate. Ignoring correlations and some other statistical niceties, the first order effect of testing N images at each recognition attempt is to change the performance of the system to $FMR(N) \sim N^{*}FMR(1)$ and $FNMR(N) \sim FNMR(1)^{N}$. The important point is that FMR(N) *increases* slowly (linearly) with N, whereas the FNMR(N) *decreases* much more rapidly (exponentially).

Current implementations of iris2pi can generally provide a better FMR than is necessary in most recognition applications. Hence, a slightly higher FMR(N) can be a trade-off for a much lower FNMR(N). This enables us to use lower quality recognition images. Lower quality on the recognition image can then be converted into larger standoffs and larger capture volumes for a fixed system cost. IOM systems can be designed to operate at the bottom end of the image quality standards set by the ISO. In particular the systems can be designed for ~100 pixels across the iris and somewhat less than 40 dB SNR.

The essay now considers each of the system parameters in turn and discusses the relevant tradeoffs.

Capture Volume

Capture volume is the product of two factors: field of view (an area) and ▶ depth of field (a distance). The field of view is product of the width and height of the region in focus for the imager. Field of view is determined by the pixel count of the camera sensor. A 2048 × 2048 pixel camera will give a field of view of 20.48 × 20.48 cm at a subject iris resolution of 100 pixels/cm. IOM systems achieve a large field of view by accepting an iris resolution of only 100 pixels/cm and by using cameras with high pixel counts.

Depth of field is the distance along the axis connecting the subject and the camera over which the iris is in focus – without additional adjustment of the lens. Depth of field is a characteristic of the lens system and is discussed in detail elsewhere [11]. Depth of field increases with increasing F# of the lens, which in turn reduces the amount of light that gets through the lens and the SNR of the image. IOM systems achieve effective depth of fields of the order of 10 cm at standoffs of 2–3 m by providing sufficiently bright illumination to enable use of higher F#'s while still maintaining a good signal-to-noise-ratio.

Standoff

Once the pixel size of the imager and subject resolution are chosen, the magnification of the lens system is determined by the rules of geometrical optics. The magnification combined with the camera standoff determines the focal length of the lens through the lens equation $1/f = 1/p + 1/q$ where f is the focal length, p is the subject distance, and q is the image distance. The magnification links p and q: $M = q/p$.

IOM systems achieve camera standoffs of 2 m or more by using long focal length lenses (~200 mm) with F#'s determined by ▶ diffraction limits and depth of field considerations.

Iris recognition systems almost always use some form of active near IR illumination. Both, the illumination and the camera have standoffs. IOM systems achieve illumination standoffs of 1–2 m by using arrays of powerful near IR LEDs to deliver irradiance of the order of 2 mW/cm^2 to the subject. The irradiance at the subject is limited by safety considerations. The ▶ threshold limit values (TLV) for near IR irradiation are published by the ACGIH [12]. The TLVs limit both the subject irradiance (W/cm^2) and the source radiance (W/cm^2-sr), which is a characteristic of the LEDs used in the illuminator.

Residence Time

The product of residence time and the image capture rate (frame rate) of the camera gives the number of images captured during a recognition attempt. This must exceed one – preferably more than one to get advantage from the multiple attempt tradeoff described above.

IOM system use relatively high frame rate cameras (15–60 fps) to minimize residence time while still getting enough images to reliably perform recognition.

Subject Motion

Subject motion can be divided into two types: longitudinal motion, along the camera axis, and transverse motion, perpendicular to the camera axis. Subject motion has two primary effects on image capture: (1) it limits the residence time within the capture volume and (2) it introduces motion blur, which effectively degrades the image resolution.

If the capture volume is 10 cm thick – set by the depth of field – and the subject is walking through the volume with a longitudinal velocity of 1 m/s, he will transit the volume in 0.1 s; his residence time is 0.1 s. The product of his residence time and the image capture rate must be greater than 1, as noted above.

A rough measure of motion blur is the amount of motion, in pixels, that occurs during the acquisition of an image. To maintain the resolution of the system this needs to be less than 1 pixel. Longitudinal velocity, v_L has a blurring effect that is the result of magnification change as the subject approaches the camera. The size of the change is approximately the ratio of the distance moved to the camera standoff. Since the iris is ~100 pixels across, the magnification change must be kept below ~0.01. If the shutter time of the camera is t, and the camera standoff is d_c, then $0.01 > v_L\, t\, /d_c$. For a 1 ms shutter, and 2 m standoff, the longitudinal velocity must be less than 20 m/s – about twice as fast at the world record (~10 s) in the 100 m dash. Hence, the maximum longitudinal velocity is set by required residence time, rather than motion blur.

Transverse velocity is a much bigger issue. Motion perpendicular to the camera subject axis simply moves the pixels at the iris as much as the motion. For a transverse velocity v_T, a shutter time, t, and a subject resolution, δ_s (m/pixel) $v_T\, t < \delta_s$ is required. For 100 pixels/cm and a shutter time of 1 ms, $v_T < 0.1$ m/s – much smaller than the longitudinal velocity limit.

The IOM systems use shuttered cameras and strobed illumination to freeze the subject motion. However, 1 ms is, at present, a practical limit on the strobes and shutters. The IOM systems use human factors engineering to minimize the transverse motion of the subject as they move toward the camera. The subject is given a target to walk towards and is instructed to walk in a straight line.

The strobe and shutter control the amount of light that reaches the camera sensor. Shorter strobes/shutters at the same peak irradiance reduce the number of photons reaching the sensor and reduce the SNR. It is possible to increase the light intensity during the strobe to recover part of this, but there are limitations imposed by the maximum current that can be used to drive the LEDs and maximum safe radiance and irradiance set by the TLVs.

Ambient Light

Ambient light can interfere with iris recognition systems in various ways. Bright ambient lights can cause extreme pupil constriction and can cause subjects to squint; both of these effects can cause difficulty in generating a good quality iris template from the image. However, short of reducing the ambient, there is little that can be done to ameliorate these physiological responses. Bright ambient lights can also cause specularities on the iris that interfere with the generation of iris templates. This factor can be ameliorated by using narrow band illumination with narrow band filtering on the camera. It can also be ameliorated with the strobed illumination and shuttered camera used for blur reduction.

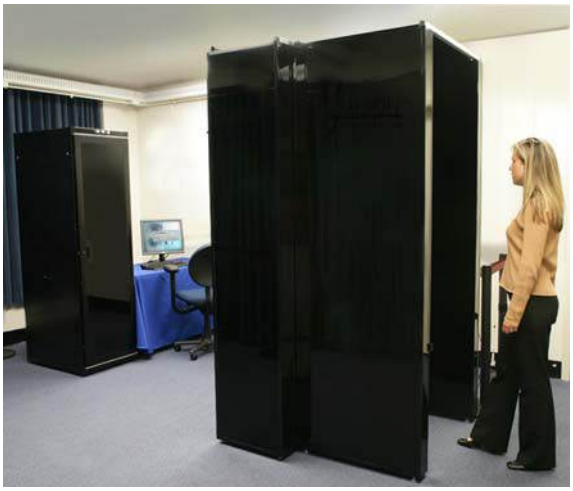
IOM systems reduce the effect of ambient illumination by using optical filters that block a large fraction of the ambient light while passing the active illumination and by using strobed illumination and opening the shutter only during the strobe, thereby rejecting the ambient light that impinges on the subject during the time the strobe is off.

Ease of Use

The IOM systems use human factors designed to maximize ease of use. People find it relatively easy to walk down the center of a portal, to walk along a line painted on the floor, or to walk directly toward a target from a given starting point. The fields of view of the IOM systems have been chosen to insure that a person carrying out such a task will, as a matter of course, pass their eyes through the capture volume of the system.

The IOM systems have been designed to take advantage of the motion of the subject toward the camera, to avoid the need for the subjects to position themselves at the focus of the system – they will walk through it naturally as shown in Fig. 2. The instructions for a subject about to use an IOM portal for the first time are:

1. Open your eyes
2. Look straight ahead at your reflection in the camera cover
3. Walk down the center of the portal at a moderate pace
4. Try to be recognized



Iris on the Move™. Figure 2 An Iris on the Move™ portal system. (Photo provided courtesy of Sarnoff Corporation.)

System Approach

In the world of software engineering, a "spaghetti design" is anathema – modularity is the mantra – with each module distinct and independent. In the IOM systems, photons thread essentially all of the hardware components. Modularity, in one sense, is not possible. Changes in the camera sensor will have unavoidable repercussions for the illumination. The key to designing an IOM system is to recognize the need for system wide optimization.

Future Work

Researchers at numerous institutions are working to expand capture volumes, extend standoff distances, reduce residence times and decrease the cost of iris recognition systems. They are also working to reduce constraints on subject pose, a topic not covered here. The development of higher resolution camera sensors with higher sensitivity and higher frame rates and the development of higher output near IR LEDs will enable these improvements.

Alternative iris recognition algorithms are also being developed, as noted above. It is conceivable that a new algorithm may enable interesting new tradeoffs in the design of iris cameras that will help us further extend the ease of use of iris recognition.

Related Entries

- ▶ Biometric Data Interchange Format, Standardization
- ▶ Biometric Sensor and Device, Overview
- ▶ Biometric System Design
- ▶ Iris Acquisition Device
- ▶ Iris Encoding and Recognition
- ▶ Iris Image Quality
- ▶ Iris Recognition, Overview
- ▶ Iris Super-Resolution

References

1. Matey, J.R., Naroditsky, O., Hanna, K., Kolczynski, R., LoIacono, D., Mangru, S., Tinker, M., Zappia, T., Zhao, W.Y.: Iris on the Move™: acquisition of images for iris recognition in less constrained environments. Proc. IEEE **94**(11), 1936–1947 (2006).
2. Mansfield, A., Kelly, G., Chandler, D., Kane, J.: "Biometric product testing: final report" (CESG Contract X92A/4009309), Centre for Mathematics and Scientific Computing, UK National Physical Laboratory (2001)
3. Daugman, J.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 21–30 (2004)
4. International Biometrics Group: Independent testing of iris recognition technology, Final report, May 2005, NBCHC030114/0002. Study commissioned by the US Department of Homeland Security
5. Bertillon, A.: La couleur de L'Iris. Annales de Demographie Internationale **7**, 226–246 (1886)
6. Schwartzman, J. producer.: Never Say Never Again [motion picture]. Warner Bros. (1983)
7. Daugman, J.: "Biometric personal identification system based on iris analysis." US Patent 5,291,560, 1 Mar 1994
8. Phillips, P.J., Scruggs, W.T., O'Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M.: "FRVT 2006 and ICE 2006 large-scale results." NISTIR 7408, March 2007. See also http://iris.nist.gov/
9. Daugman, J.: Probing the uniqueness and randomness of IrisCodes: results from 200 billion iris pair comparisons. Proc. IEEE **94**(11), 1927–1935 (2006)
10. ISO/IEC: "Biometric data interchange formats – Part 6: Iris image data" ISO/IEC JTC 1/SC 37 N 504 (2004)
11. Smith, W.J.: Modern Optical Engineering. McGraw-Hill, New York (2000). ISBN 0-07-136360-2. http://en.wikipedia.org/wiki/Depth_of_field
12. ACGIH: TLVs and BEIs (2006). ISBN 1-882417-62-3. www.acgih.org

Iris Quality Metrics

- ▶ Iris Image Quality

Iris Reader

- ▶ Iris Acquisition Device
- ▶ Iris Device

Iris Recognition, Overview

Yung-Hui Li[1], Marios Savvides[2]
[1]Language Technology Institute, Carnegie Mellon University, Pittsburgh, PA, USA
[2]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Synonym

Iris Biometric

Definition

Iris recognition emerges as one of the most useful modalities for biometrics recognition in last few decades. The goal of iris recognition is to recognize human identity through the textural characteristics of one's iris muscular patterns. The procedures for iris recognition usually consist of four stages: image acquisition, iris segmentation, feature extraction, and pattern matching. The iris recognition has been acknowledged as one of the most accurate biometric modalities because of its high recognition rate. It has been applied in the field of border control and national security. More and more countries and private companies have shown interests to use the technique of iris recognition. Large scale application of iris recognition in daily life is just a matter of time.

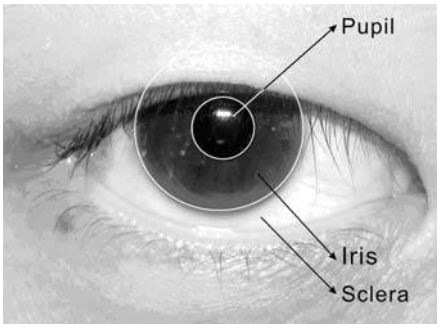
Introduction

The goal of biometric recognition is to recognize human identity by comparing the features of their physiological or behavioral characteristics. There are dozens of such characteristics that can be found and used for such purpose. Among them, fingerprint, face, and voice are most familiar to most people. However,

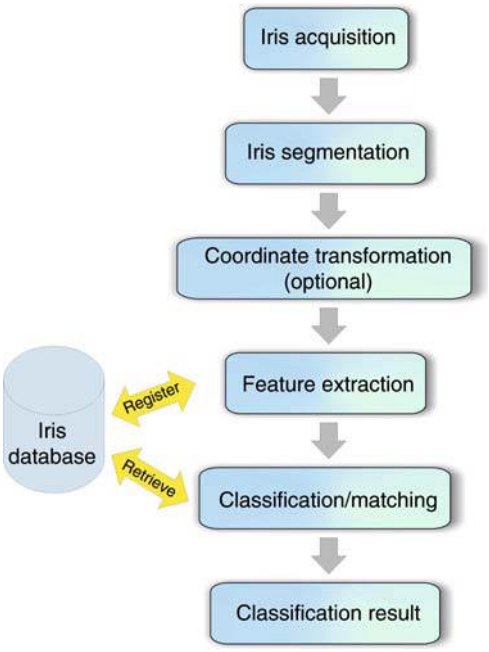
the performance (in terms of the recognition rates) for all the above modalities is not good enough for the purpose of applications that require high level of security.

Not very long ago, identity recognition based on iris texture patterns is proposed as another modality for biometric recognition [1–3]. Iris is the area that appears as an annular shape between the pupil and the sclera, as illustrated in Fig. 1. There are many advantages of using iris as the modality of biometric recognition. For example, iris pattern is developed and becoming stable after the eighth month of gestation. Besides, iris is well-protected by cornea, not easily changed by external factors. Therefore, iris patterns are very stable, relative to other modalities such as face or voice, and very suitable to be used as a unique feature for identity recognition. Another advantage is that randomness of the muscle of the iris region makes possibility of sharing exactly the same iris pattern among different persons very low, which provides high discriminability for this feature.

A typical iris recognition system consists of the following stages: iris image acquisition, iris segmentation, feature extraction, and pattern matching, as illustrated in Fig. 2. Image acquisition is of vital importance in iris recognition performance because the better the image quality, the more accurate result the system can achieve. After iris image is acquired, the next step is to locate the position of the iris and pupil. Several techniques have been proposed and most of them are based on image gradient and edge detection. After iris has been located, features of the iris textures can be extracted. What feature to extract and how to extract them is also important issue in iris recognition. Good features give high inter-class variation and low within-class variation, which is desired in common biometric recognition system. Finally, one should be able to compare features from different irises and obtain a score of similarity which tells how similar these two irises are (or a score of distance, which tells how much different they are). Decision of classifying test image as authentic or imposter is made by comparing the similarity score with a given threshold.



Iris Recognition, Overview. **Figure 1** Illustration of the areas of pupil, iris and sclera.



Iris Recognition, Overview. **Figure 2** Flow chart of a typical iris recognition system.

Iris Image Acquisition

The area of iris is very small compared with the area of the whole face. Therefore, it is not a trivial task to take a clear iris image with high quality. On the other hand, user-friendliness is also an important factor which has to be considered in the process of iris image acquisition. According to the user-friendliness and the image quality, iris image acquisition device can be categorized into three different groups. They are traditional iris acquisition device, middle distance iris acquisition device, and long distance iris acquisition device. Each of them is introduced in the following paragraphs.

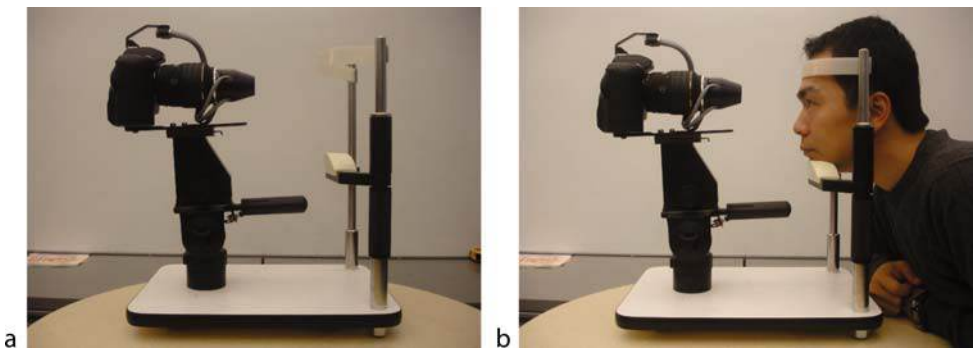
For traditional iris acquisition devices, in order to take iris images that contain enough useful detailed

information, people usually use cameras which has specialized lens set up on a fixed stand. At the same time, there is also a fixed stand for subjects to lean their heads against. In this way, the system minimized the possibility of subtle movement of subject's head during the iris image acquisition stage, and maximized the quality of the acquired iris image. Exemplar images of traditional iris acquisition device are shown in Fig. 3.

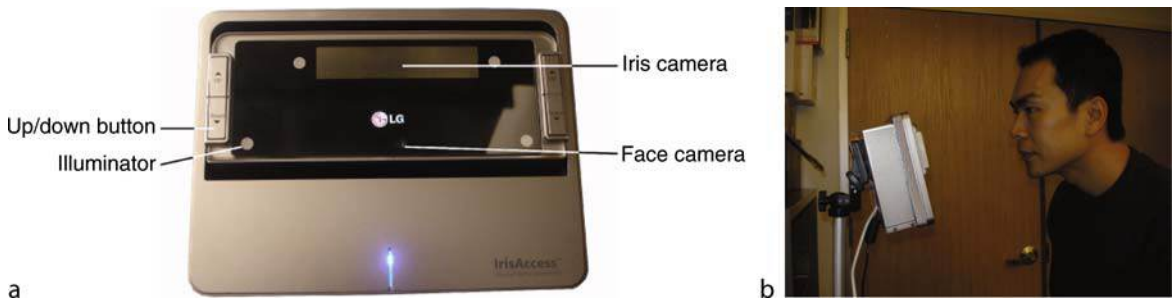
Traditional iris acquisition devices are good for capturing high quality iris image, but extremely inflexible and not user-friendly. It requires high level of user-cooperation to successfully acquire iris image. Is it possible to lower the level of required user-cooperation and at the same time, maintain the high quality of the eye image? Recently, progress of innovation in iris acquisition device has made it possible to lessen the requirement of user-cooperation while maintaining high quality iris images. They can be categorized as middle distance iris acquisition devices because they can capture users' iris images even when users stand at a distance of 50–100 cm away from the camera. Users are not required to put their heads against a rack. Therefore, it is more user-friendly and convenient for general public.

One of examples of middle distance iris acquisition devices is the LG iCAM4000. The functional unit of LG iCAM4000 is shown in Fig. 4, and an example picture is shown when it is used to acquire a user's iris image.

Long distance iris acquisition devices are able to capture iris images even when users are standing at a further distance, or even when they are moving. One example of this kind of devices is the iris-on-the-move™ (IOM) manufactured by Sarnoff Corporation. The system provides a portal which subjects are expected to walk through. At the end of the walking aisle, four high resolution infrared (IR) cameras are placed inside a cabinet. During the time subjects are walking in the portal, IR illuminators located at the wall of the portal radiate IR light, at the same time, high resolution cameras which tuned to sync with the illuminators take continuous shots of the subjects. Then a face and eye detector is applied on the image to find the location of the eyes. The whole iris acquisition process is fully automated and it only requires subjects to walk through the portal, without standing still or placing



Iris Recognition, Overview. Figure 3 (**a**) Example of traditional iris acquisition device (**b**) Example picture to illustrate how to use it.



Iris Recognition, Overview. Figure 4 (**a**) LG iCAM4000 iris camera (**b**) Example picture to show how to use it.

their heads at a fixed position. In this way, the system alleviates the inconvenience of the traditional iris acquisition devices while being able to maintain a reasonable quality of the eye pictures.

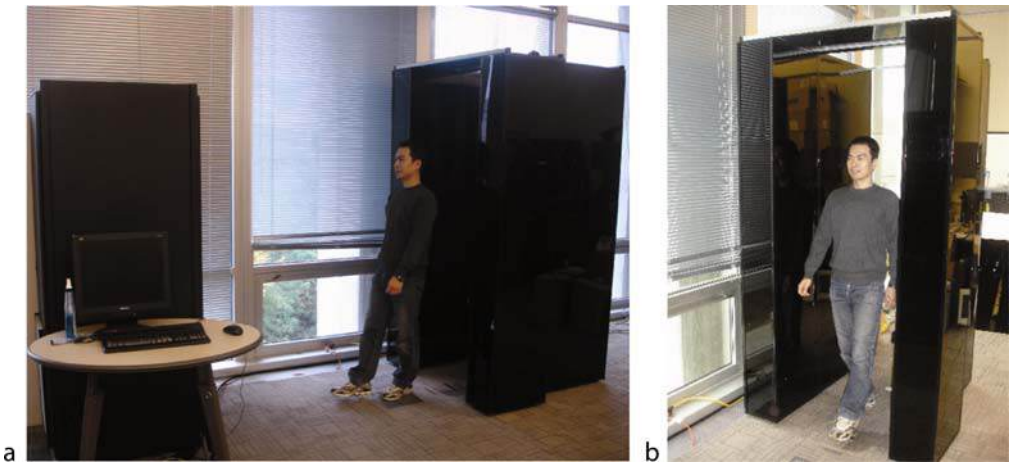
Figure 5 shows pictures of IOM system.

It is very intuitive that the system which requires least user cooperation (more user-friendly) has the highest probability of getting eye images of lowest quality. For example, the IOM system is very user-friendly. It does not require users to stand still at a particular point to get their iris images. However, the eye images taken from IOM are very easy to be blurred or out of focus. This is because the images may be taken when the subjects are walking in a region which is outside of the depth of the field of the cameras.

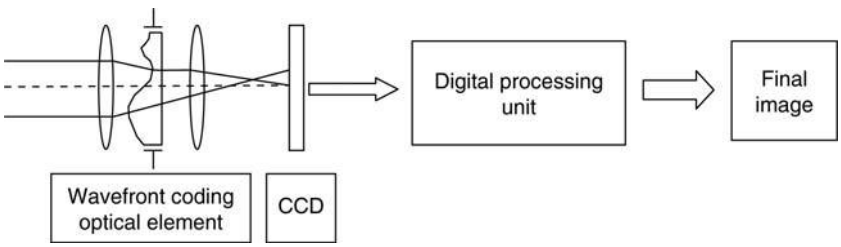
Recently, a new camera system design has been proposed to address this issue. Dowski and Johnson proposed a low cost imaging system which combines the nonrotationally symmetric aspheric optical elements and digital signal processing in a fundamental manner to vastly extend the depth of field of the imaging systems [4]. It is called "Wavefront Coding" imaging system. The block diagram of the Wavefront Coding system is shown in Fig. 6. The optical section is modified by adding a generalized aspheric Wavefront Coding optical element near the aperture stop. This will make the images formed on CCD to have a specially well-defined blur that is insensitive to misfocus. Digital image processing technique is applied to the blurred image to reproduce the sharp and clear image. Optical systems aided with Wavefront Coding are able to reproduce a clear and focused image even when the subject is standing at the out of focus zone. For more details, please refer to [4].

Iris Segmentation

No matter which kind of iris acquisition device is used, the eye image taken usually looks like the one in Fig. 1. Therefore, it is necessary to locate the iris region in this image and focus our future process



Iris Recognition, Overview. Figure 5 (**a**) IOM system. Camera cabinet is on the left; the portal is on the right, behind the subject. (**b**) Example picture of subject walking through the portal.



Iris Recognition, Overview. Figure 6 Block diagram of Wavefront Coding imaging system.

(feature extraction and pattern matching) on this region alone. The process of locating iris region is called iris segmentation.

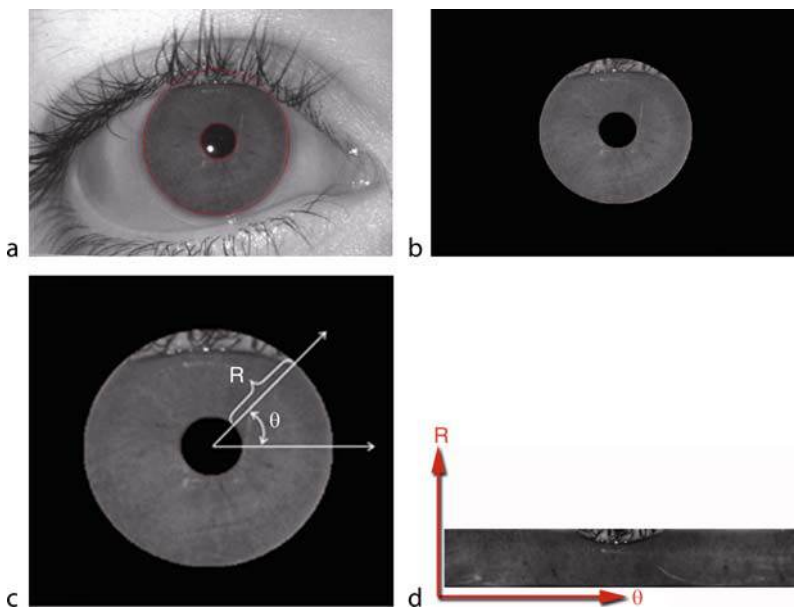
As stated above, iris region is of annular shape. Therefore, it is intuitive to segment iris with two circles: one circle indicates the boundary between iris and pupil, and the other indicates the boundary between iris and sclera. The texture in the region between these two circles is exactly what we need for further process.

There exist different algorithms which serve the same goal of iris segmentation. One of the most popular algorithms is proposed by Wildes [5]. There are two steps in this algorithm. The first step is to apply a gradient-based edge detector on the whole eye image to generate an edge-map. The edge-map tells us the position where strong edges exist (strong differences in pixel values). Intuitively, those positions are possible candidates of the iris boundaries since the two boundaries of iris, both inner and outer, are the positions where high pixel contrast takes place. The second step is to find the exact two boundaries from this edge-map. The method Wildes used is a voting scheme. Every circle on a 2D plane can be described by three parameters, coordinate in x and y axis, and the radius r. Therefore, one can construct a three-dimensional (3D) space, where each dimension represents one parameter. Every positive pixel on the edge-map can vote to the point (x, y, r) in the 3D parameter space as long as this positive pixel is on the perimeter of the circle which is parameterized as (x, y, r). Since different circles can pass through the same point, it is possible for 1 pixel on the edge-map to vote to multiple points in 3D parameter space. At last, the point which has the highest accumulated votes represents the most likely circle in the original eye image. An example of iris segmentation is shown in Fig. 7a.

Daugman proposed another segmentation scheme [6, 7] which is different from what Wildes proposed. He suggested computing the argument which maximizes the result of an integrodifferential operator on the original image. It is described in the following equation:

$$\max_{(r,x_0,y_0)} \left| G_\sigma * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r} \mathrm{d}s \right|. \qquad (1)$$

The triple (r, x_0, y_0) which satisfied Eq. 1 is the best candidate parameter for a circle in the original image. By applying Eq. 1 twice, we will be able to locate the boundaries for both iris and pupil. To locate the eyelid boundary, one only needs to change the form of integrodifferential operator in Eq. 1 from circular to



Iris Recognition, Overview. Figure 7 (**a**) A raw eye image with iris boundaries identified as red circles (**b**) segmented iris image after masking out unuseful region (**c**) illustration of the meaning of parameter R and θ (**d**) unwrapped iris image in polar coordinate.

arcuate, with spline parameters, because the shape of eyelid is most likely to be arcuate.

After observing large number of iris images, one sooner or later discovers that in fact, the boundaries of iris and pupil are not exactly circular. They appear as circular images of low resolution. However, for images of higher resolution, it is easy to see that using a circular boundary to describe them is not very accurate.

Therefore, Daugman, in his later work [8], proposed a new segmentation method. He proposed to use a more flexible model to represent both boundaries. This method is called "active contour," or "snakes." By using snakes, the boundary of the pupil and iris is not bounded to circular. They can be of any shape. Therefore, the discovered boundaries can fit to the real data more closely, and bring performance enhancement in the pattern matching stage. The points of snakes are initialized at the points from sampling N regularly spaced angular samples of radial gradient edge data. We have to first estimate the Fourier series expansion of those samples of radial gradient edge data, as described in the following equation:

$$C_k = \sum_{\theta=0}^{N-1} r_\theta e^{-2\pi ik\theta/N}, \tag{2}$$

where $\{r_\theta\}$ is the radial gradient edge data, $\theta = 0 \sim N-1$.

We compute C_k from $k = 0$ to $k = M-1$. Then we compute an approximation of the iris boundary $\{R_\theta\}$ for $\theta = 0 \sim (N-1)$, as described below:

$$R_\theta = \frac{1}{N} \sum_{\theta=0}^{M-1} C_k e^{2\pi ik\theta/N}. \tag{3}$$

The number of M controls the smoothness of the curve. The larger the M, the more flexible the curve is; and the smaller the M, the less flexible the curve. By choosing M carefully one can make the discovered boundaries fit the data more closely and not being affected by the noises (eyelid or eyelashes).

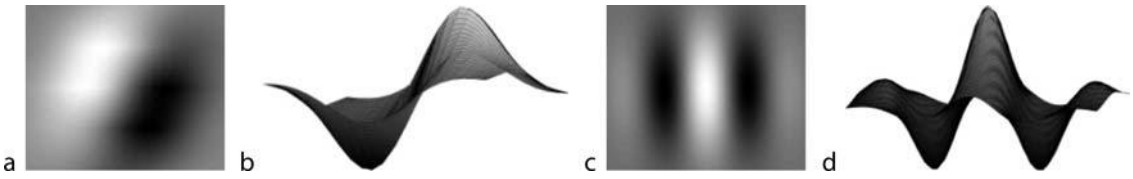
Most of the time, after iris localization, we would like to do an image coordinate transformation on the iris image. The goal of coordinate transformation is to transform the annular-shaped iris region into rectangular shape, by mapping pixel values from Cartesian coordinate to polar coordinate. The process is illustrated in Fig. 7. In Fig. 7b, the region which is unimportant for iris recognition is masked out, leaving only iris texture visible. This iris image is displayed in Cartesian coordinate. If we pick the center of the pupil as the origin, and the horizontal line as x-axis, every point on iris region can be indexed by another coordinate, which is polar coordinate. Polar coordinate is parameterized with two parameters: R and θ, and the computation of R and θ is illustrated in Fig. 7c. If we take out every pixel in iris region in Fig. 7c and replot them according to the coordinate (R, θ), we will get Fig. 7d, which is the iris texture map in polar coordinate.

There are two advantages of coordinate transformation. First, the size of the pupil area is not always the same. It contracts when the ambient lighting is strong and dilates when the ambient lighting is weak. Therefore, if we would like to take the iris image in Cartesian coordinate to perform pattern matching, one problem is the nonlinear deformation of the pattern, which require much more complex technique to solve. Instead, if we perform coordinate transformation for every iris texture, no matter how much pupil contracts or dilates, the resulting iris texture map in polar coordinate will remain the same. This saves us a lot of trouble during the later matching stage.

The second problem is, when we take pictures of eyes, if we are not using traditional iris acquisition device, it is highly possible that the position of our camera has relative rotational shift to the subject's head. It is very likely to happen because users can freely move their head in any direction, as long as the eye images can be captured. If there is indeed relative rotational shift between cameras and the eyes, this is going to cause another trouble during later matching stage. This is because matching two patterns with potential rotational shift is a difficult problem in pattern recognition field. Instead, if the iris image is unwrapped to polar coordinate before feature extraction and matching stage, rotational shift in Cartesian coordinate becomes translational shift in horizontal direction (θ axis). Therefore, all we need to do is simply shift the pattern horizontally in different step sizes and perform matching again to get correct matching result.

Feature Extraction

After iris region has been identified and cropped out, the next step is to perform feature extraction on the iris texture. The goal of feature extraction is to extract the discriminative characteristic of the iris texture and store it in a more compact way so that it is more effective to perform pattern matching in a later stage.



Iris Recognition, Overview. **Figure 8** Examples of two filters which can be used for iris feature extraction. (**a**) Directional Gaussian derivative filter, in 2D view (**b**) Directional Gaussian derivative filter, in 3D view (**c**) Gabor filter, in 2D view (**d**) Gabor filter, in 3D view.

Most of the time, in the field of image processing and pattern recognition, the feature extraction is accomplished by applying filters on the input images. In literature, the filters which have been used extensively in iris recognition, are Gabor filters, proposed by J. Daugman [2, 6, 9, 10, 7]. Gabor filters can be seen as complex sinusoids modulated by Gaussian envelope. They can be expressed as the following equation:

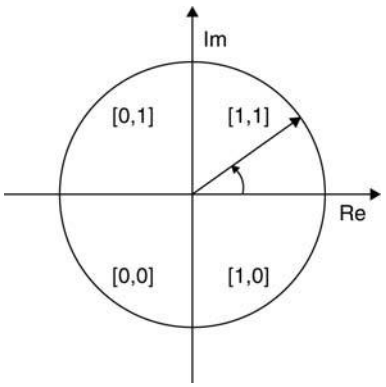
$$\Psi(x,y) = A \cdot e^{\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)-j\omega(\cos\theta)x-j\omega(\sin\theta)y\right]}, \quad (4)$$

where, by convention, the standard deviations of the envelope are inversely proportional to frequency ω:

$$\sigma_x = k_x \frac{2\pi}{\omega}$$
$$\sigma_y = k_y \frac{2\pi}{\omega}. \quad (5)$$

We can generate multiple Gabor filters of different sizes, orientation, or frequency by substituting different parameters in Eq. 4. Gabor filters of different orientation and frequency can capture different texture details in iris texture. An example of Gabor filter as well as Gaussian derivative filter are shown in Fig. 8.

We get a two-dimensional (2D) complex-valued plane after applying each Gabor filter on iris texture. Daugman proposed a ▶ phase-quadrant quantization scheme to further quantize the 2D complex-valued plane. Because amplitude information depends on extraneous factors such as imaging contrast, illumination, and camera gain, it is not very discriminating. On the contrary, phase information on a complex-valued filter response plane has been shown to have much more discriminative information in biometric recognition [7, 11]. Therefore, performing phase quantization would discard useless information and at the same time, make feature representation more compact and easy to handle.

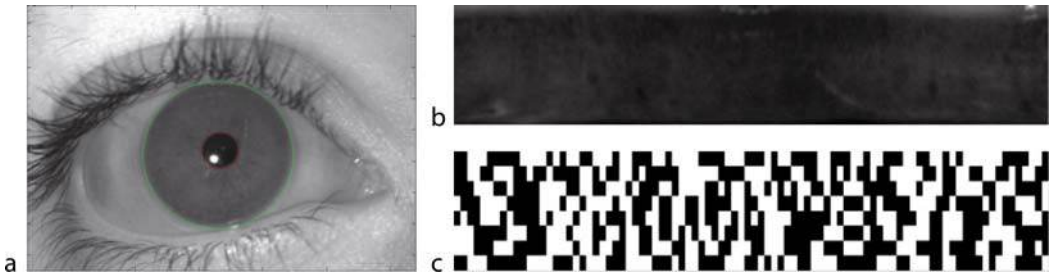


Iris Recognition, Overview. **Figure 9** Illustration of Daugman's phase-quadrant demodulation scheme. According to the quadrant each complex number falls in, the angle of the phasor is quantized to one of the four quadrants, setting two bits of phase information. This process is repeated across all the iris with many wavelet sizes, frequencies, and orientations.

The phase-quadrant quantization scheme amounts to a patch-wise phase quantization of the iris pattern, by identifying in which quadrant of the complex plane each resultant phasor lies. For every location on the complex plane, the sign of real and imaginary part jointly defines in which quadrant the phasor lies. Two bitcodes are used to encode the four possible quadrants. The phase-quadrant quantization scheme is illustrated in Fig. 9.

The phase-quadrant demodulation process is repeated across all the plane of the response of the filters, with many different wavelet sizes, frequencies, and orientations to extract totally 2048 bits. Figure 10 shows the raw eye image, unwrapped iris texture in polar coordinate, and the iris code after feature extraction.

Besides Gabor filters, there are also other type of 2D filters that can be used in iris feature extraction. Zhu



Iris Recognition, Overview. Figure 10 (**a**) Eye image with segmentation in Cartesian coordinate (**b**) unwrapped iris image in polar coordinate (**c**) iris code computed by applying Gabor filters on (**b**) then perform bit-wise quantization.

et al. proposed to use 2D wavelet transform to perform feature extraction [12]. By applying 2D wavelet transform on an iris texture map, one can get a set of sub-images of different resolution levels. They proposed to use the mean and variance (or standard deviation) of each wavelet sub-images to be the features.

Ma et al. proposed to use circular symmetric filters to perform iris feature extraction [13, 14]. The kernels of circular symmetric filters can be defined as follows:

$$G(x,y,f) = \frac{1}{2\pi\delta_x\delta_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\delta_x^2}+\frac{y^2}{\delta_y^2}\right)\right]M(x,y,f)$$
$$M(x,y,f) = \cos\left[2\pi f\left(\sqrt{x^2+y^2}\right)\right], \quad (6)$$

where f is the frequency of the modulating sinusoidal function, δ_x and δ_y are the space constants of the Gaussian envelope along the x and y axis, respectively. The difference between circular symmetric filters and Gabor filters is that Gabor filter of a fixed parameter focus only on information obtained from a specific orientation; while circular symmetric filters extract information from all orientation. Circular symmetric filters can also put more emphasis on a particular orientation, by changing the parameter δ_x and δ_y. After iris texture has been convolved with circular symmetric filters, they proposed to extract the mean and the average absolute deviation of the magnitude of local patches (of size 8×8) of the convolved image as feature vectors. Finally, Fisher linear Discriminant analysis was applied to reduce the dimensionality of the feature vectors.

One last thing to note is that, very often, iris texture is occluded by other objects, for example, eye lashes, eye lids, or specular reflection from eye glasses, it is not that every point on the iris texture map is useful for pattern matching. Therefore, it is necessary to compute an iris mask, which is exactly the same size as iris texture, to indicate which part of the map is really iris texture, and which part is not. This mask will be used in the stage of pattern matching, as described in the next section.

Pattern Matching

After iris code and its mask have been computed, the next step is to match two irises to see if they are coming from the same class. Obviously, to complete the task, the only thing one has to do is to count how many bits of these two irises share the same value. One thing to note is that when we compare the differences of each bit, we must not count the bit which belongs to the region that is not iris texture. Therefore, we have to mask out non-iris region before we compute the matching score.

Another thing to note is that since we already quantize the feature of iris texture into binary values, there are only two possible values for each location on iris map (1 or 0). There exists one operation which can be executed extremely fast to compare if 2 bits share the same value, which is the exclusive-or operation (XOR). XOR returns zeros if both bits are the same (both of them are zeros or ones), ones if they are different (one of them is 0 and the other is 1). Therefore, by performing bit-wise XOR operation, one can compute the distance score for two irises in extremely high efficiency.

To summarize, suppose we want to compare the distance between iris A and B, whose two phase code bit vectors are denoted {*code_A*, *code_B*} and whose

mask bit vectors are denoted {*mask_A, mask_B*}, it can be computed simply by following equation:

$$HD = \frac{\|(code_A \otimes code_B) \cap mask_A \cap mask_B\|}{\|mask_A \cap mask_B\|}, \quad (7)$$

where $\otimes$ denotes bit-wise_ XOR operation, $\cap$ denotes bit-wise AND operation, and $\|\cdot\|$ denotes the norms of the resultant bit.

The computed score is a metric of "distance," indicating how different these two irises are. It is called "Hamming Distance" (HD). As one can see from Eq. 7, Hamming Distance is normalized by the norm of the effective region. Therefore, it will always have values between 0 and 1, where 0 indicates if iris A and B are exactly the same and 1 indicates if they are completely opposite.

Ideally, if two irises are from the same class, the Hamming Distance between them should be close to 0. On the other hand, if two irises are from different classes, due to the property of statistical independence, the probability for each bit of one iris to match the same bit of another iris should be 50%. Therefore, the expected Hamming Distance for two irises which come from different classes should be 0.5. In practice, the values of Hamming Distance will not be exactly 0 or 0.5. The distribution of the value of Hamming Distance for authentic comparison will be a Gaussian distribution, centered at 0, and the distribution of the value of Hamming Distance for imposter comparison will be another Gaussian distribution, centered at 0.5. If the quality of the input iris image is high enough, in most cases, these two Gaussians would not intersect with each other, or they only intersect with each other in tiny portion. Therefore, a proper threshold can be chosen to minimize both the ► False Positive Rate and ► False Negative Rate.

Summary

Iris recognition is an emerging field for biometric recognition. Substantial research efforts have been involved in this field to push the performance of iris recognition to the limit. Literature has shown that it is one of the biometric modality that has high performance, high universality, high distinctiveness, high permanence, and low chances of circumvention [15]. As the technological innovation of iris acquisition keeps advancing, it is becoming more user-friendly and more popular. As the computational power of hardware grows exponentially and the size of chip keeps decreasing, iris segmentation, feature extraction, and matching can all be executed faster in much smaller devices. In the near future, it is very promising that iris recognition system will be widely accepted, not only in the application at national security level, but also in private companies, public services, and private residency.

Related Entries

► Biometrics Overview
► Biometric System Design
► Iris Acquisition Device
► Iris Encoding and Recognition

References

1. Flom, et al.: Iris recognition system. US Patent 4641349, 3 Feb 1987
2. Daugman, J.: Biometric personal identification system based on iris analysis. US Patent 5291560, 1 Mar 1994
3. Wildes, et al.: Automated non-invasive iris recognition system and method. US Patent 5572596, 5 Nov 1996
4. Dowski, E.R. Jr., Johnson, G.E.: Wavefront coding: a modern method of achieving high-performance and/or low-cost imaging systems. Proc. SPIE 3779, 137 (1999), DOI:10.1117/12.368203
5. Wildes, R.: Iris recognition: An emerging biometric technology. Proc. IEEE **85**, 1348–1363 (1997)
6. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. **15**(11), 1148–1161 (1993)
7. Daugman, J.: How iris recognition works. IEEE Trans. Circ. Syst. video Technol. **14**(1), 21–30 (2004)
8. Daugman, J.: New methods in iris recognition. IEEE Trans. Syst. Man Cybern. Part B, **37**(5), 1167–1175 (2007)
9. Daugman, J.: Statistical richness of visual phase information: Update on recognizing persons by iris patterns. Int. J. Comput. Vis. **45**(1), 25–38 (2001)
10. Daugman, J.: The importance of being random: Statistical principles of iris recognition. Pattern Recognit. **36**(2), 279–291 (2003)
11. Savvides, M., Kumar, B.V.K.V., Khosla, P.K.: Eigenphases vs eigenfaces. In: Proceedings of the 17th International Conference on Pattern Recognition 2004 (ICPR 2004), vol. 3, pp. 810–813 vol. 3, 23–26 (Aug 2004)
12. Zhu, Y., Tan, T., Whang, Y.: Biometric personal identification based on iris patterns. In: Proceedings of International Conference on Pattern Recognition, pp. 2801–2804. Barcelona, Spain (2000)

13. Ma, L., Tan, T., Wang, Y., Zhang, D.: Personal identification based on iris texture analysis. IEEE Trans. Pattern Anal. Mach. Intell. **25**(12), 1519–1533 (2003)
14. Ma, L., Wang, Y., Tan, T.: Iris recognition using circular symmetric filters. In: Proceedings of International Conference on Pattern Recognition, pp. 414–417 (2002)
15. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circ. Syst. Video Technol. **14**(1), 4–20 (2004)

Iris Recognition Algorithms

An iris recognition algorithm is a method of matching an iris image to a collection of iris images that exist in a database. There are many iris recognition algorithms that employ different mathematical ways to perform recognition. Breakthrough work by John Daugman led to the most popular algorithm based on Gabor wavelets.

► Iris Acquisition Device

Iris Recognition at Airports and Border-Crossings

John Daugman
Computer Laboratory University of Cambridge, Cambridge, UK

Synonyms

CANPASS; CLEAR; Iris recognition immigration system (IRIS); NEXUS; Privium; RAIC

Definition

As case illustrations of generic biometric applications, there are at least five different modes in which automated personal identification by iris recognition is used at airports: (1) international arriving passengers can clear Immigration control at iris-automated gates without passport or other identity assertion if they have been enrolled in a preapproved iris database; (2) departing passengers can receive expedited security screening and check-in as low-risk travelers if enrolled in an iris database following background checks; (3) airline crew members use iris recognition for controlled access to the secure air-side; (4) airport employees gain access to restricted areas within airports such as maintenance facilities, baggage handling, and the tarmac; and (5) arriving passengers may be screened against a watch-list database recording the irises of persons deemed dangerous, or of expellees excluded from entering a country. All such existing programs use the Daugman algorithms for iris encoding and recognition because of the need to process iris images fully at the speed of the video frame rate (30 frames/s) and to search databases at speeds of about a million IrisCodes per second, and the need for robustness against making False Matches in large database searches despite so many opportunities. However, the threat models posed for the different applications are distinctive, depending on whether an attacker's goal is a False non-Match (a concealment attack, e.g., in a watch-list application) or a False Match (an impersonation attack, e.g., to be taken for a registered traveler in an expedited Immigration control or trusted-traveler deployment). Likewise, the business models vary for these different uses, depending on whether the traveler pays for the convenience of expedited processing, or an airport owner pays for the facility's enhanced security and productivity, or a government funds such a technology deployment both to improve process efficiency and to achieve national security goals.

Introduction

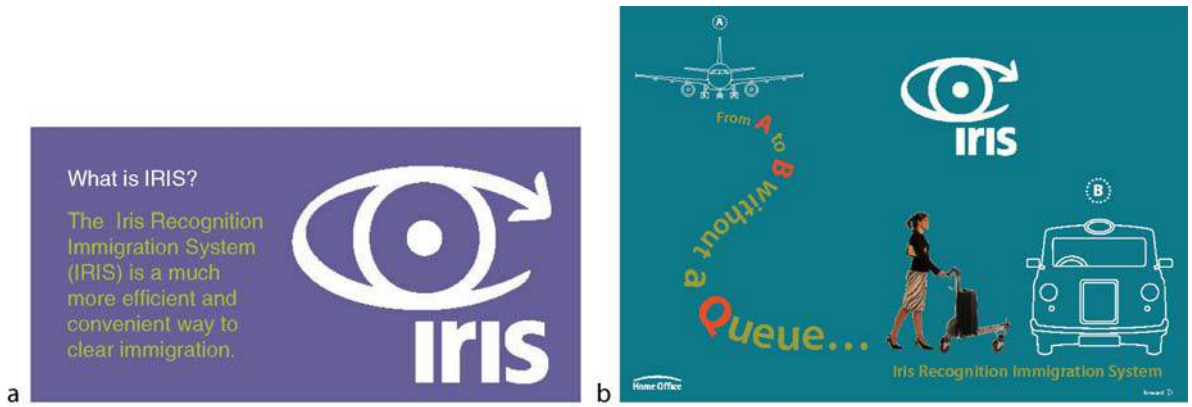
Most deployments of biometric systems have as their main purpose either enhancing the security and reliability, or enhancing the convenience and efficiency, of an identification process. In some applications either security or efficiency dominates the requirement, while the other is less important. For example, in identifying theme park visitors biometrically in lieu of ticketing, efficiency is much more important than security; whereas for biometric applications within prisons or detention centers, just the opposite is the case. In airports, however, both of these objectives are paramount, and neither can be compromised. Excelling simultaneously at both

objectives creates special challenges for biometric systems, because the design strategies and indeed the core technology choices that may maximize throughput volumes are not necessarily the same as those that maximize identification accuracy. This article reviews five ways in which automated iris recognition [1] is used within airports and at border-crossings, with special attention to those trade-offs and design issues.

Arriving International Passengers: Iris Recognition Instead of Passport Presentation at Immigration Control

The use of biometrics as living passports, removing the need for actual passport presentation at Immigration control, was pioneered in the UK in 2002. A 6-month trial of the *EyeTicket JetStream* system allowed a total of 2,000 frequent travelers from North American to London Heathrow Airport to enroll their ▶ IrisCodes and thereby to bypass Immigration control upon arrival, passing instead through an automated iris recognition gate. The trial was deemed fully successful and led eventually to a large-scale system deployed by the UK Home Office, called IRIS: *Iris Recognition Immigration System* [2]. Based on the same core Daugman algorithms [1] but with a more user-friendly interface, the *IRIS* system is today deployed at most major UK airports, including all five terminals at Heathrow. The architecture incorporates a centralized database of enrolled IrisCodes so that travelers can use the system regardless of their airport or terminal, although this also makes the system vulnerable to interruptions in communications links or reductions in bandwidth. Such network failures in the first year of deployment occasionally interrupted the service. Nonetheless, as of May 2008, the UK Border Agency announced that more than one million passengers had successfully used the system, with enrollments increasing by about 2,000 per week, and with the system handling about 15,000 arrivals per week. By substituting for passport presentation, the system replaces long queues at arrivals with an expedited automated clearance at iris camera gates within a matter of seconds (Fig. 1).

A crucial aspect of the *IRIS* system is that it operates in *identification mode* to determine a passenger's identity, not in a mere *verification mode* in which an identity is first asserted (for example by presenting a token, passport, or smartcard) that is then simply verified. The requirements of biometric operation in identification mode by exhaustively searching a large database are vastly more demanding than one-to-one verification mode in which only a single yes/no comparison with one nominated template is required. If P_1 is the False Match probability for single one-to-one verification trials, then $(1-P_1)$ is the probability of not making a False Match in single comparisons. The likelihood of successfully avoiding any in each of N independent attempts is therefore $(1-P_1)^N$, and so P_N,



Iris Recognition at Airports and Border-Crossings. **Figure 1** The UK Government's *IRIS* program has enabled more than a million registered travelers to enter the country via several British airports using only automatic iris recognition for identification, in lieu of passport presentation or any other means of asserting an identity.

the probability of making at least one False Match when searching a database containing N different patterns, is

$$P_N = 1 - (1 - P_1)^N \quad (1)$$

Observing the approximation that $P_N \approx NP_1$ for small $P_1 << \frac{1}{N} << 1$, when searching a database of size N an identifier needs to be roughly N times better than a verifier to achieve comparable odds against making False Matches. In effect, as the database grows larger and larger, the chance probability of making a False Match also grows almost in proportion. Obviously the frequency of False Matches over time also increases with the frequency of independent searches that are conducted against the database. These considerations make it vital that such identification applications operating by exhaustive search use a biometric modality and algorithms that generate score distributions with extremely rapidly attenuating tails, when different persons are compared. (These issues are discussed and documented in more detail in the article *Score Normalization Rules in Iris Recognition.*) In the absence of such rapidly attenuating distribution tails, the system would drown in False Matches when the search databases become large. In this connection, it is noteworthy that in the UK where the *IRIS* program optionally replaces passport presentation, the Border Control Development and Strategy Group forecasts that by 2015, the number of international passengers entering the UK annually will exceed 150 million.

Several other countries are also deploying the same iris recognition algorithms as a substitute for passport presentation. One of these is The Netherlands, where iris-based border-crossing has been used since 2003 for frequent travelers into Amsterdam Schiphol Airport; members of the *Privium* program pay an annual fee to be able to use automated iris gates for clearing Immigration, in lieu of waiting in queues for passport presentation. Another country with a similar but larger deployment is Canada, where the *CANPASS* program operates in all the eight international airports (Edmonton, Winnipeg, Calgary, Halifax, Ottawa, Montreal, Toronto, and Vancouver) with about 40 kiosks at each [3]. Both US and Canadian citizens or permanent residents are entitled to enroll in this iris-based system for entering Canada. In addition, the *NEXUS* program operated jointly by the USA and Canada allows border-crossing in both directions across their shared border using iris recognition for preapproved travelers (Fig. 2).

Finally, motorcyclists who commute daily across the border between Malaysia and Singapore for work use iris recognition to avoid the long queues for checking passports and ID papers. The Singapore *Iris Border*

Iris Recognition at Airports and Border-Crossings. **Figure 2** At Schiphol Airport (Amsterdam NL), the *Privium* Program has a membership of about 40,000 frequent travelers. They pay an annual fee to use the iris recognition system at automated gates, thereby avoiding the queues at Immigration for passport presentation.

Control for Motorcycles allows 3,000 commuters to cross the border efficiently using "registered iris" lanes with automated gates, as may be seen in an on-line video [4]. The motorcyclists in these lanes remain on their bikes; the gate is equipped from the side with iris cameras, including one for a passenger on the bike. Riders must stop and remove their helmets, but they do not assert their identity. Rather, identification is performed by exhaustive search of the enrolled iris database linked to the fully automated gates. The system also maintains a watch-list that is checked.

Departing Passengers: Expedited Check-in, Security Screening, and Border Controls

The US Transportation Security Administration (TSA) and Department of Homeland Security (DHS) in 2005 began a public/private partnership known as the *Registered Traveler* (RT) Program to make airport security procedures more efficient for departing passengers deemed to be trusted. Under this program, dozens of US airports have deployed iris and fingerprint recognition systems to confirm the identity of "trusted travelers" who have been vetted by the TSA and approved for expedited security screening. Bypassing the long lines that have become a feature of airport security checkpoints since September 2001 is a benefit for frequent travelers, who pay an annual fee of about $100 for this privilege. It is also an enhancement for TSA security processes which can become more focused and can take advantage of the background vetting that was done when a person was enrolled in the scheme by virtue of being deemed a minimum security risk.

Although baggage X-ray and metal detection checks remain universal, enrollees in these systems face less intrusive screening (e.g., they can keep their coats and shoes on and laptops in their bags), and they enjoy access to a reserved fastlane with shorter delays. These privileges are asserted by presenting a smartcard credential that contains their biometric data as well as other information, all under two layers of encryption and readable only by TSA card-readers. Biometric kiosks in the departure fastlanes read the cards and confirm passengers' identity with iris cameras or fingerprint readers. The network is interoperable across some 30 US airports, and the list is steadily expanding [5]. Beginning with Orlando Airport in July 2005, some of the major participating US airports today include JFK, LaGuardia, Newark, Dulles, Regan, Denver, and San Francisco International Airports. The largest such program is called *CLEAR*, operated by Verified Identity Pass, which had 175,000 enrolled members as of July 2008 [5]. Additional newer participants in the *Registered Traveler* public/private partnership with the TSA include FLO, Unisys, and Vigilant.

In Europe, for travelers who are nationals of the 25 EU countries that have entered into the Schengen Agreement for harmonized border control, the identification formalities for crossing into and out of the Schengen Zone are done by iris recognition at kiosks in certain airports. The first such deployment was at Frankfurt/Main Airport and is known as the Automated and Biometrics-based Border Checks (*ABG*) initiative. This multinational project is led by Germany's Federal Ministry of the Interior and Federal Border Police. The stated objectives of the scheme are to eliminate the use of fraudulent travel documents and multiple identities, to speed trusted travelers across borders, and to allow greater productivity for border officials.

Iris recognition is also used for other, nonsecurity related enhancements for departing passengers at airports such as Milan's Malpensa and Tokyo's Narita Airport. Under the *Simplifying Passenger Travel* scheme implemented by the Ministry of Justice in Japan, the JAL Group offers streamlined procedures for passenger check-in and boarding pass issuance, as well as immigration control at departure and certain "*e-airport*" utilities and facilities. These services are provided at iris-enabled automated kiosks and gates in departure areas, as illustrated in Fig. 3.

Airport Employees: Access Control to the Tarmac, Aircraft, and Restricted Areas

Probably the most traditional use of biometric recognition is for physical access control, to ensure that only authorized persons enter restricted facilities. This classical mode of biometric deployment is found at many airports, controlling access to aircraft maintenance facilities, baggage handling areas, the tarmac and other secure zones.

The Canadian Air Transport Security Authority uses iris recognition to verify the identities of airport

Iris Recognition at Airports and Border-Crossings. **Figure 3** In the *e-airport* deployment at Tokyo Narita Airport, iris recognition is used for expedited check-in of departing passengers. In dozens of US airports, *Registered Travelers* approved by the Transportation Security Administration receive expedited security screening once their identities are proved by fingerprint or iris recognition.

workers at all the 29 major airports in Canada. Iris biometric data are embedded within an ID card called *RAIC: Restricted Area Identification Card*. Workers must present this card and verify their identities at iris cameras controlling automated portals. Similar systems are deployed at Schiphol Airport (Amsterdam) for 30,000 airport employees; at Albany Airport for baggage handlers; and at New York JFK Airport for access to the tarmac at two terminals. Some airports such as Douglas International (Charlotte) have also deployed iris recognition gates specifically for pilots and other airline crew members to reach airside more efficiently.

Finally, it is noteworthy that an International Standard specifically related to biometric identification of airport employees was published in 2008. The ISO/IEC 24713-2 Standard gives normative requirements on *Biometric Profiles for Interoperability and Data Interchange: Physical Access Control for Employees at Airports* [6]. The scope of this Standard includes recommended practices for enrollment, watch-list screening, prevention of duplicate token issuance, and employee identity verification. It also describes architectures and business processes appropriate to token-based identity management within the secure environment of an airport.

Watch-list Screening of Arriving Travelers

The rapid search capabilities of iris recognition, and its robustness against making False Matches despite the fact that large search databases create many opportunities for such errors, have led to the deployment of this technology for watch-list screening. The largest such deployment is in the United Arab Emirates, where visa-bearing travelers arriving at any of the 32 air, land, and sea ports of entry are processed with iris recognition cameras as illustrated in Fig. 4.

Known as the *Expellee Tracking and Border Security Iris System*, the scheme was launched in 2001 by the UAE Ministry of Interior. A noteworthy aspect of the UAE is that among its 5.4 million residents, about 85% are foreign nationals [7] on work permits. Because of this large foreign labor force drawn by economic opportunities much better than elsewhere in the Middle East and South Asia, men outnumber women by a factor of 2.74 among persons in the 15–65 age group [7], and the border-crossing volume of migrant workers whose homeland roots are elsewhere is very high (some 12,000 per day). In 2001 an amnesty was granted to all foreign nationals who had overstayed

Iris Recognition at Airports and Border-Crossings. **Figure 4** In the United Arab Emirates deployment of iris recognition at all the 32 air, land, and sea ports, travelers are screened against a watch-list of expellees, or persons deemed to be a security risk, before being allowed to enter the Emirates.

their work permits or committed other visa violations, but a condition of the amnesty waiver of penalties was that such persons were expelled from the Emirates for some period of time, and their iris patterns were registered in a database. This action enabled enforcement of the ban on re-entry and defeated thousands of attempts to return under false identities and with fake travel documents. Over the period 2001–2007 the database of expellees' IrisCodes was enlarged with IrisCode databases of foreign nationals who had been imprisoned for crimes such as prostitution or drugs trafficking, and of persons deemed to be security risks or unwelcome for other reasons.

Today this iris watch-list contains 1.2 million IrisCodes from persons of 156 nationalities. All travelers seeking visa entry into the UAE via any port have their iris images acquired by cameras as shown in Fig. 4, so that their IrisCodes can be computed and matched exhaustively against the full database. Since on average some 12,000 such persons arrive at the UAE each day, about 14 billion IrisCode comparisons are performed daily across a dedicated network. The *IrisFarm* architecture is a distributed host/client system with a single central database maintained by the Abu Dhabi Police, linked over a network of communication channels to clients that send IrisCode queries to it from all ports of entry. The average turn-around time is about 2 s. Because every query IrisCode is compared exhaustively with all on the watch-list, the total volume of such iris comparisons performed over the years of operation now number in many trillions [8]. Tens of thousands of persons have been caught trying to re-enter the UAE under false identities, who are turned away but who often make repeated attempts, and the UAE Ministry of Interior hails the system as a huge success. The system is now expanding into neighboring Gulf States including Jordan and Oman, and it will be linked with an iris-based national identity and border-crossing system being procured in the Kingdom of Saudi Arabia.

System Design Contrasts and Vulnerabilities

The most important differences among the various systems reviewed in this article are (1) whether they operate in *identification mode,* in which no identity is asserted but identity is determined by searching a database, versus *verification mode* in which a token like a smartcard is used to assert a particular identity that is then simply verified one-to-one; and (2) whether the objectives of a valid user or an attacker are to be matched to an identity on a database, or not.

Identification is vastly more demanding than one-to-one verification, both in terms of search space and

comparison speeds, and in terms of the requirement to avoid any False Matches despite what may be a huge number of opportunities to make them if the database is large. If a weak biometric system such as face recognition is used, an attacker would have an excellent chance to be matched just by chance against at least one person in a trusted traveler database, if that were the only test and if the database were larger than a few hundred or perhaps a thousand. For this reason, weaker biometrics rely on smartcards or other tokens to assert a particular identity, so that only one comparison must be executed successfully. But presentation of a token makes the process more cumbersome, and in any case it has no value for watch-list screening.

Nearly all deployments of iris recognition operate in identification mode by exhaustive search of a database, because the technology's speed and accuracy allow it. The exceptions to this mode are (1) the *Privium* system because Dutch law forbids the storage by the State of personal data like biometrics, and so the citizens alone retain it; and (2) the *CLEAR* program because a smartcard is used for several other purposes in the transaction anyhow. In both of these cases the use of a storage token makes it unnecessary to perform identification by searching a database.

In identification systems operating by database search, it is necessary to combat the inevitable net increase in the likelihood of chance False Matches as the size of the search databases grow. This form of probability summation is the same phenomenon as arises when playing the game of Russian Roulette an increasing number of times. In the case of the iris recognition algorithms [1, 8] used in all current iris deployments, combatting this is accomplished by minute adjustments in the decision threshold with search database growth, keeping the net False Match probability minuscule. Further details about these processes are given in the accompanying article, *Score Normalization Rules in Iris Recognition.*

In a trusted traveler scheme (*CLEAR*, *IRIS*, *Privium*, etc.), the objective of an attacker is to impersonate another person – either a particular person, or anyone at random just by accident – who is registered in the trusted database. The likelihood of success by blind chance (a "zero effort attack") is minuscule in the case of iris, but much higher if a printed contact lens can be produced to mimic a particular target individual's iris. In a watch-list deployment such as the UAE one, the objective of an attacker is simply to look like anybody other than himself (or anyone else registered in the watch-list). Such a "concealment attack" by means of printed contact lenses is easier than an "impersonation attack," and indeed it can even be attempted simply by being uncooperative. Therefore, these border security systems incorporate tests for the vitality, or "liveness," of iris patterns including their motion and deformation with changes in the pupil size, which obviously does not occur if printed on a contact lens. Similarly, the standard algorithms perform biometric quality assessments to detect extremely dilated pupils or excessively closed eyelids, as indicators of possible attacks. However, the struggle between countermeasure and new counter-countermeasure continues and escalates relentlessly.

Related Entries

► Iris Encoding and Recognition Using Gabor Wavelets
► Score Normalization Rules in Iris Recognition

References

1. Daugman, J.G.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**, 21–30 (2004)
2. UK Border Agency, Project IRIS website. http://www.ukba.homeoffice.gov.uk/managingborders/technology/iris/
3. Canadian Border Services Agency, CANPASS website. http://cbsa-asfc.gc.ca/publications/pub/bsf5017-eng.html
4. Singapore Iris-Based Border Control for Motorcyclists. http://www.youtube.com/watch?v=HieaASl9sE8&feature=related
5. CLEAR Verified Identity Pass website. http://www.flyclear.com
6. International Organisation for Standards: Biometric Profiles for Interoperability and Data Interchange, Part 2: Physical Access Control for Employees at Airports. ISO/IEC **24713-2** (2008)
7. Demographics of UAE, 2008. http://en.wikipedia.org/wiki/United_Arab_Emirates
8. Daugman, J.G.: Probing the uniqueness and randomness of IrisCodes: results from 200 billion iris pair comparisons. Proc. IEEE **94**, 1927–1935 (2006)

Iris Recognition Immigration System (IRIS)

► Iris Recognition at Airports and Border-Crossings

Iris Recognition Operational Range

The maximal transversal region along the optical axis of an imaging system within which an iris recognition algorithm provides accurate identification. Typically, iris recognition operational range is significantly affected by the system's depth of field.

► Wavefront Coded® Iris Biometric Systems

Iris Recognition Performance Under Extreme Image Compression

John Daugman, Cathryn Downing
Computer Laboratory, Cambridge University, Cambridge, UK

Definition

The compressibility of images is usually gauged by their subjective appearance and by metrics for the amount of distortion that can be tolerated. In the context of biometrics, compressibility can be gauged objectively by measuring the impact of compression schemes on recognition performance compared to baseline performance. Standard biometric methodologies such as Receiver Operating Characteristic (ROC) curves are perfectly suited for measuring the impact of compression on performance. It is possible for performance actually to benefit from slight image compression, as has been seen both with fingerprint and iris recognition, because high frequency noise is the first thing lost; but at more severe levels, compression must become detrimental. For iris recognition, it is possible to compress images to as little as 2,000 bytes through a combination of methods including cropping, region-of-interest (ROI) isolation and JPEG2000 wavelet coding, while suffering only a little reduction in recognition performance. This is important because Governments and Standards organizations prefer that biometric data be stored in a relatively raw, unprocessed form in order to remain algorithm-neutral and future-proof. It is also mathematically important because of insights from information theory and complexity theory related to minimal description length, entropy, compressibility, and discriminability.

Introduction

A watershed event in biometric informatics occurred in 1993 when the FBI digitized a vast library of fingerprint cards that had been stored in acres of filing cabinets and adopted the Wavelet Scalar Quantization (WSQ) protocol [1] to compress these images, achieving compression ratios in the range of 10:1 or 15:1 without detectable loss of detail. Compressibility is a fundamental issue for biometrics, not only for mundane practical reasons related to storage requirements and data transmission times, but also for abstract mathematical reasons related to information content and pattern recognition. Governments, regulatory bodies, and international standards organizations often mandate the storage of relatively raw data rather than processed biometric templates, hoping thereby to preserve interoperability and to keep biometric data "vendor-neutral and future-proof" while the algorithms for pattern description and recognition inevitably evolve and improve. But raw images as data objects are almost a thousand times larger than the biometric templates ultimately computed from them. Hence, the conflicting goals of reducing the size yet preserving the information content of biometric data make it important to understand how various possible compression schemes impact on biometric recognition performance. This article studies these questions in the context of iris recognition, and reviews data [2] showing that it is possible to compress iris images to within about a factor of two of the standard iris biometric template sizes with almost no impact on recognition performance.

Information Theory and Data Compressibility

Data compression is one of several disciplines rooted in information theory having relevance to biometric technologies for identifying persons, and its significance extends beyond the practical matters of storage requirements and data transmission times. One of Shannon's fundamental insights in formulating

information theory [3] was that the entropy of a random variable measures simultaneously its information content (expressed in bits) and its compressibility without loss (to the same number of bits). This link between entropy, informativeness, and compressibility extends also to other measures that apply to biometrics. For example, the relative entropy between two distributions is one way to measure how well a biometric technique separates samples from same versus different persons. The amount of variability in a given biometric system across a population, or in different samples from the same source, is also captured by conditional entropies, with larger entropy signifying greater randomness. Finally, the similarity between pairs of biometric templates is reflected by their mutual information: the extent to which knowledge of one sample predicts the other. All of these properties are deeply connected with the compressibility of biometric data.

An extreme variant of Shannon's insight was expressed by Kolmogorov [4] in his notion of minimal description length, which defined the complexity of a string of data as the length of the shortest binary program that could generate the data. Creating that program "compresses" the data; executing that program "decompresses" (generates) the data. Fractal image compression is based on this idea; and a data string is said to be Kolmogorov incompressible if the shortest program that can generate it is essentially a data statement containing it, so the data are then their own shortest possible description. Within biometrics, this notion has appeared implicitly under a different rubric in work on *synthetic biometrics,* seeking methods for artificially synthesizing a biometric image that is indistinguishable in practice from some actual biometric image. Pioneering work in this direction was done by Terzopoulos and Waters [5] for facial images and sequences, by Cappelli et al. [6] for fingerprints, and by Cui et al. [7] and by Zuo et al. [8] for iris images. In future, such programs for generating particular biometric images might therefore serve as ways to "compress" them in Kolmogorov's sense; and one might even anticipate biometric recognition by comparison of the synthesizing programs. The present article explores a combination of image compression methods applied to iris images, specifically probing the question of how aggressively they can be compressed without impairing iris recognition. A convergence between compressed image size and biometric description length begins to emerge.

Schemes for Iris Image Compression

Iris templates (e.g., ▶ IrisCodes) are usually computed from a polar or pseudo-polar coordinate mapping of the iris, after locating its inner and outer boundaries [9, 10, 11]. However, if simple circular boundary models are imposed, polar mappings depend strongly upon the choice of origin of coordinates which may be prone to error, uncertainty, or inconsistency, especially since the true iris boundaries are often not actually circular. Unlike rectilinear coordinates, for which a shift error has no more effect than a shift, in polar coordinate mappings a shift error in the choice of coordinate origin can cause large distortions in the mapped data, with no way to recover from the deformed sampling.

A pioneering study of iris compressibility was undertaken by Rakshit and Monro [11] showing that if segmented and normalized iris data were extracted in polar form, this "unwrapped" polar data structure could be compressed to 2,560 bytes or even less without impairing recognition performance. However, because this approach stores the iris image in polar form, it is not robust against errors in assigning the origin of coordinates or the loss of iris data when circular borders are inappropriately enforced. It also suffers from making the accurate detection of eyelid boundaries for exclusion of eyelid regions very difficult, since the available pixel data is tightly cropped around the iris and ignores valuable shape-constraining data from the boundary between sclera and eyelid. Finally, the polar unwrapping has the consequence of highly nonuniform sampling, with pixels near the outer perimeter of the iris sampled with a density 2.5 times lower than those near the pupil. Preservation of the original rectilinear format of an iris image is more veridical than polar unwrapping methods because pixels retain constant size and spacing.

Avoiding the problems introduced by polar unwrappings of iris images, the authors [2] investigated three compression schemes that retain the native rectilinear image array format but compress it to as little as 2,000 bytes while still allowing very good recognition performance on the difficult Iris Challenge Evaluation (ICE-1) iris database available from the National Institute of Standards and Technology (NIST) [12] which includes many poor quality images. Interoperability was documented between those images when subjected to the three compression regimes and their uncompressed form, and it was found that on average only

2–3% of the bits within the computed IrisCodes are affected even when the net image data reduction factor reaches 150:1.

Tools for Iris Image Compression: JPEG, JPEG2000, Region-of-Interest Extraction

Clearly, a first step in image data reduction is to crop the iris images from the standard format of 640 × 480 pixels with 8 bits grayscale data per pixel, consuming 307,200 bytes, to a smaller region of 320 × 320 pixels centred on the iris. This was done by running the eye-finding part of the standard algorithms [10] that are used in all current public deployments of iris recognition, on all images in the publicly available NIST [12] ICE1Exp1 database, which contains 1,425 iris images from 124 Subjects with "ground-truth" information given about which images were taken from the same iris. The algorithms correctly localized the iris in all images and produced from each one a new cropped image of 320 × 320 pixels with the iris centred in it. For those NIST images in which the iris was partly outside of the original image frame, the missing pixels were automatically replaced with black ones. For those in which the algorithms detected that the gaze was directed away from the camera, as gauged by projective deformation of the eye shape, a corrective affine transformation was automatically applied which effectively "rotated" the eye in its socket into orthographic perspective on-axis with the camera. The first column of Fig. 1 shows three examples of iris images cropped as described earlier.

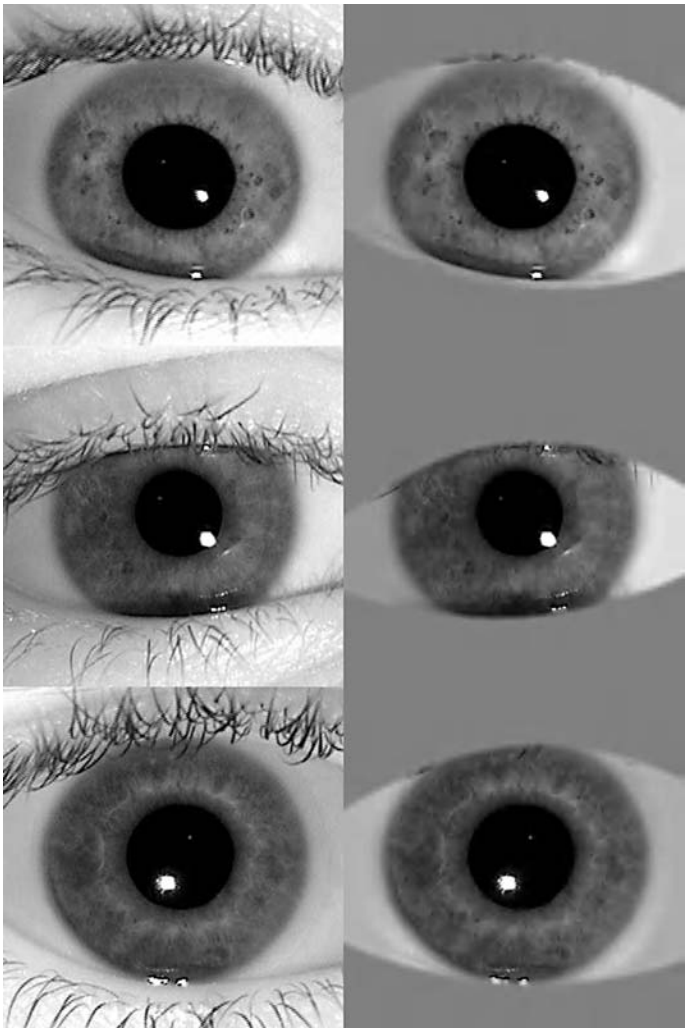
This new gallery of simply cropped images was subjected to three different compression schemes: (1) ► JPEG compression with quality factors (QF) of 70, 30, and 20; (2) JPEG compression with the same QFs but after ► Region-of-Interest (ROI) segmentation; and (3) ► JPEG2000 compression after ROI segmentation with compression factors (CF) of 20, 50, and 60, as illustrated in the second column of Fig. 1 for the case of CF = 50.

The use of cropping and JPEG compression alone (with QFs of 70, 30, and 20) produced image file sizes averaging 12,400, 5,700, and 4,200 bytes, respectively, but with large variability around these means. Including the initial threefold reduction in file size due merely to cropping the images to 320 × 320 pixels, these net data reduction factors relative to the original full-size images therefore average 25:1, 54:1, and 72:1, respectively. But further significant reductions in image data size can be achieved through the use of ROI segmentation of the iris image.

The standard lossy JPEG coding scheme [13, 14] effectively allocates bytes on an "as needed" basis, meaning that the cost of encoding uniform regions of an image is almost nil, whereas image areas containing busy textures such as eyelashes may consume much of the available information budget. In uniform regions, the only nonzero DCT (discrete cosine transform) coefficient in each block of 64 frequency components that encode an 8 × 8 pixel block (a *data unit*) is the DC coefficient specifying their average gray value; all other coefficients are 0 if the data unit is a truly uniform region, or else become 0 after coarse quantization, and so their cost in the zeroes run-length coding stage is essentially nil. Therefore JPEG encoding of iris images can be made much more efficient if all noniris parts of the image are replaced with a uniform gray value. This was accomplished for the image gallery automatically using the standard algorithms [10] for eyelids detection and fitting, and iris boundary localization, as seen in the second column of Fig. 1.

JPEG coding schemes lend themselves well to ROI differential assignment of the coding budget. Indeed the JPEG2000 standard [15, 16, 17] and even the Part 3 extension of the old JPEG standard [13, 14], support *variable quantization* for explicitly specifying different quality levels for different image regions. In JPEG2000, the MAXSHIFT tool allows specification of an ROI of arbitrary shape. This utility was explored for biometric face recognition by Hsu and Griffin [18] who demonstrated that recognition performance was degraded by no more than 2% for file sizes compressed to the range of 10,000–20,000 bytes with ROI control.

In the approach to ROI segmentation presented here, noniris regions are encoded in a way that distinguishes sclera from eyelids or eyelashes regions, so that postcompression algorithms can still determine both types of iris boundaries. Therefore two different substitution gray levels are used: a darker one signifying eyelids, and a brighter one for the sclera, computed as an average of actual sclera pixels and blending into actual sclera pixels near the iris outer boundary. Since the substitution gray levels are uniform, their coding cost is minimal and could be further reduced by using larger data units. JPEG compression of such ROI segmented iris images typically yields a further twofold reduction in file size for each of the QFs studied, while



Iris Recognition Performance Under Extreme Image Compression. **Figure 1** Cropping of iris images (*first column*), followed by region-of-interest isolation of the iris (*second column*) to achieve greater compressibility while retaining a rectilinear image array format. Substitution of noniris regions by uniform gray levels prevents wasting wavelet coding budgets on costly irrelevant structures such as eyelashes. All images in the *second column* have been JPEG2000 compressed to a data size of only 2,000 bytes.

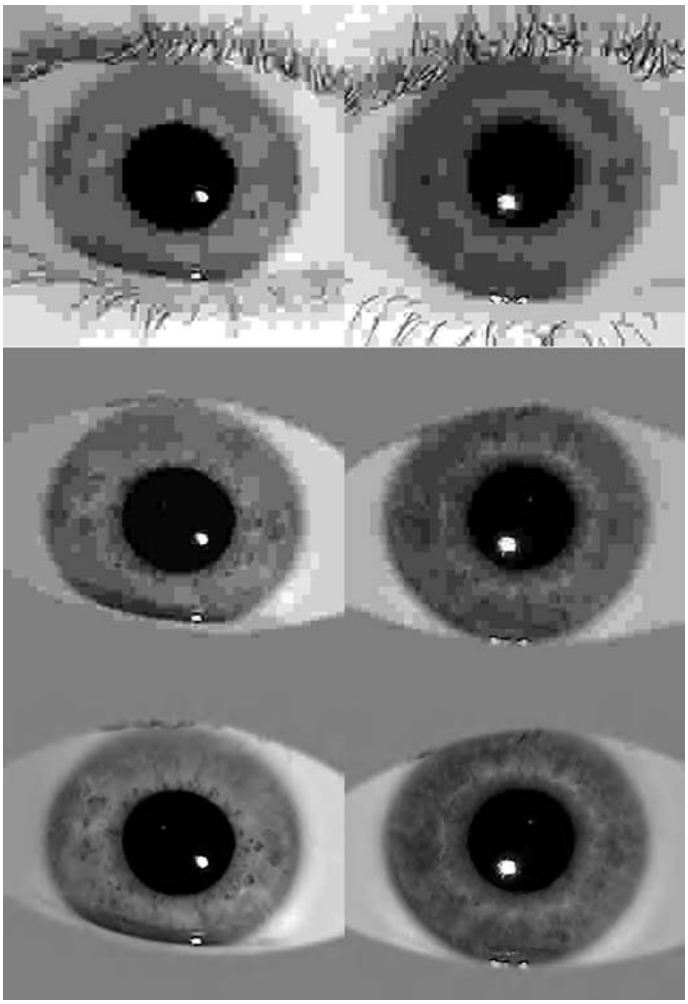
maintaining a simple rectilinear image format and easy localization of eyelid boundaries in later stages.

In 2000 a more powerful version of JPEG coding offering more flexible modes of use, and achieving typically a further 20–30% compression at any given image quality, was adopted as the JPEG2000 Standard [16, 17]. Mathematically based on a Discrete Wavelet Transform (DWT) onto Daubechies wavelets rather than the Discrete Cosine Transform (DCT), JPEG2000 does not suffer as badly from the block quantization artifacts that bedevil JPEG at low bitrates, which are due to the fact that the DCT simply chops cosine waves inside box windows with obvious truncation consequences when they are sparse and incomplete. Moreover, the different levels within the multiresolution DWT wavelet decomposition allow local areas within each image data unit to be encoded using different subbands of coefficients [17] as needed. The net superiority of JPEG2000 over JPEG in terms of image quality is especially pronounced at very low bitrates corresponding to severe compression, as investigated here, in the range of 0.15 bits/pixel (bpp).

Several mechanisms exist within JPEG2000 for heterogeneous allocation of the coding budget, including tile (data unit) definition, code-block selection allowing different DWT resolution levels in different tiles,

and DWT coefficient scaling. In the work presented here those explicit control mechanisms such as the MAXSHIFT tool were not used, but rather the same pixel substitution method as described earlier for ROI was used, for comparison purposes. Three JPEG2000 compression factors (CF) of 20, 50, and 60 were chosen, which yielded file sizes of 5,100, 2,000, and 1,700 bytes, respectively. The three images in the second column of Fig. 1 were created with a JPEG2000 CF of 50 and thus have a file size of only about 2,000 bytes. Whereas JPEG generates widely varying file sizes to deliver any given QF, JPEG2000 creates file sizes that are closely predictable from the specified CF. In the authors' experience of compressing several thousand iris images with JPEG2000, the standard deviation of the distribution of resulting file sizes was usually only about 1.6% of the mean [2], for any given CF. Predictable file size is an important benefit for fixed payload applications [19].

Finally, it is interesting to compare visually some examples of the iris images after compression to a constant data size of 2,000 bytes using each of the three different schemes for image compression that have been discussed here. In Fig. 2, each column is from the same NIST iris image; the rows represent the different schemes. The top row is a simple JPEG compression of a cropped (320 × 320) image but without ROI isolation. Most of the 2,000 byte budget

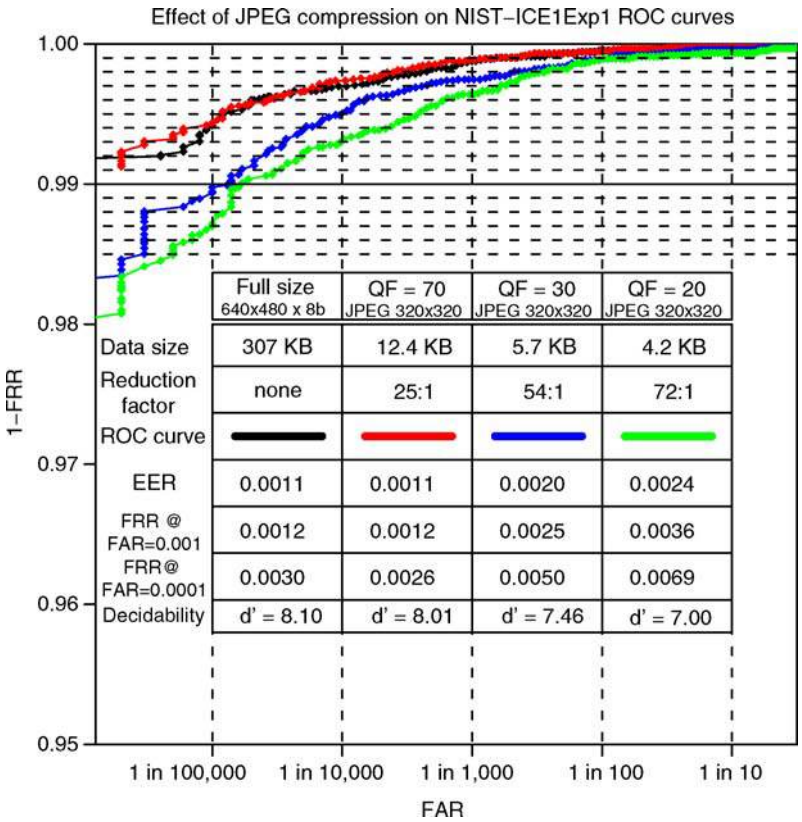


Iris Recognition Performance Under Extreme Image Compression. **Figure 2** Visual comparison of three schemes for iris image compression, for images all severely compressed to the same data size of 2,000 bytes. *First column* uses NIST image 239230; *second column* uses NIST image 239343. *Top row*: simple JPEG compression of the cropped (320 × 320) images. *Middle row*: JPEG compression of the cropped images after ROI isolation. *Bottom row*: JPEG2000 compression of the cropped and ROI-isolated images. At severe compression levels, JPEG2000 is vastly superior to JPEG.

is wasted trying to encode eyelashes, and the cost on iris texture is horrendous. The middle row shows improvement after ROI isolation, so most of the JPEG budget is allocated to the iris, but the result is still very poor. The bottom row shows the result of combining the cropping, ROI isolation, and JPEG2000 compression for the same iris images. The improvement is visually remarkable, and it is confirmed by very good iris recognition performance as summarized by the purple ROC curve (CF = 50) in Fig. 5 in the next section.

Tools for Evaluating the Effects of Iris Image Compression

Biometric recognition performance is usually measured by generating Receiver Operating Characteristic (ROC) curves, which plot the trade-off between two error rates (False Accept and False Reject Rates, FAR and FRR) as the decision threshold for similarity scores is varied from conservative to liberal. It is common to tabulate specific points on such trade-off curves, such as the FRR when the decision threshold causes an FAR of 1 in 1,000 or of 1 in 10,000, and the point at which the two error rates are equal, FRR = FAR = EER, the Equal Error Rate. Such ROC curves and tabulations are presented in Fig. 3 for the NIST [12] ICE-1 gallery, both for baseline performance (uncompressed and uncropped: black curve), and for the three QF quality factors (coloured curves) used with the simplest JPEG compression approach described earlier. The coordinates of the ROC curves are semi-logarithmic: the ordinate plots 1-FRR linearly, over just the upper 5% of its possible range, while the abscissa logarithmically spans many factors of 10 in FAR, to nearly as low as 1 in a million. The number of images and the mix of Subjects in this NIST iris



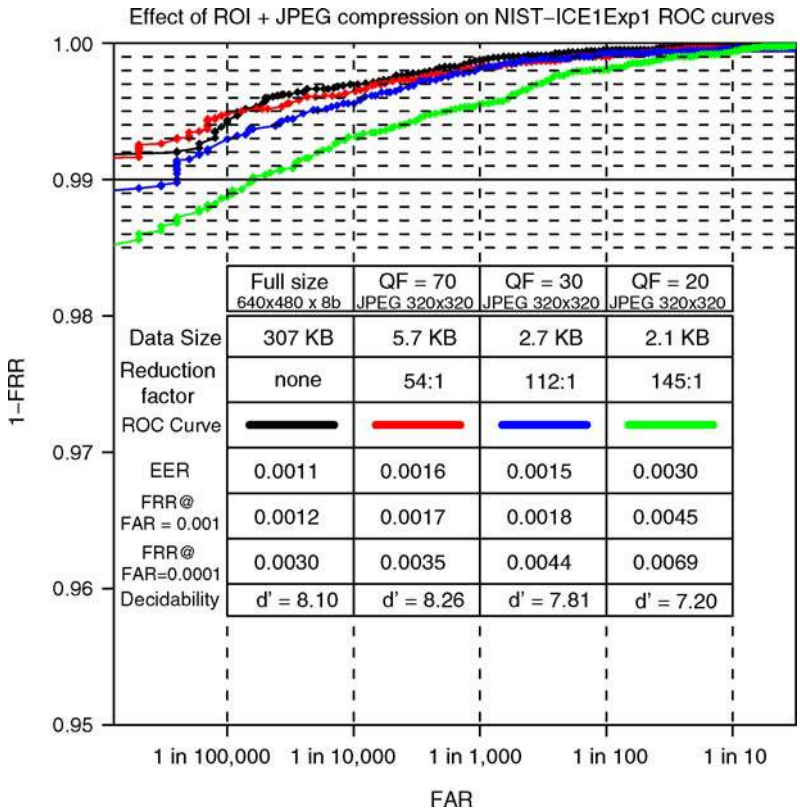
Iris Recognition Performance Under Extreme Image Compression. **Figure 3** ROC curves in semi-logarithmic coordinates for the NIST [12] ICE1Exp1 iris database, showing the impact of simple data reduction methods on performance. *Black curve* shows baseline performance on the original database of full-size images. *Red curve* shows the effect of simple cropping to 320 × 320 pixels after automatically locating and centering each iris, followed by JPEG compression at QF = 70. Blue and *green curves* show the effects of more severe JPEG compression at QF = 30 and QF = 20.

database allows 12,214 same eye matches to be tested, and it allows 1,002,386 different eye comparisons to be done, which means that one cannot measure a False Match Rate (or FAR) between 0 and 1 in a million; this determines the limit of the ROC curves on the left extreme of these graphs.

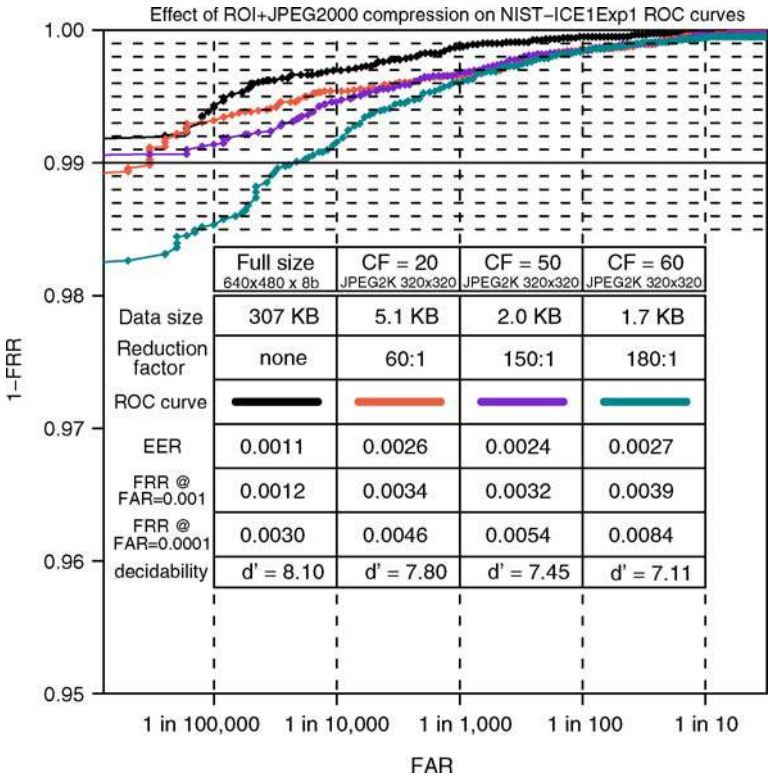
The red ROC curve in Fig. 3 shows that at a JPEG quality factor of 70 and an overall data reduction factor of 25:1, no performance loss relative to the baseline (black) ROC curve is detectable. (Indeed there is even some suggestion of a small benefit from compression, possibly due to de-noising.) The blue and green ROC curves show that for this scheme based only on image cropping and JPEG compression, using a QF in the range of 20–30 produces image file sizes in the range of 5,000 bytes but at the cost of roughly doubling the FRRs and EER compared to the error rates for uncompressed images.

The additional impact of the ROI isolation (JPEG + ROI) on iris recognition performance is gauged by the ROC curves in Fig. 4. These show that for each QF studied, iris recognition performance remained about the same as before the ROI isolation (Fig. 3), but with the achievement of a further twofold reduction in image data size, even down to the range of just 2,000–3,000 bytes per image.

Figure 5 presents the ROC curves generated by the JPEG2000 + ROI compression scheme, together for comparison with the black ROC curve for the baseline gallery (uncropped, uncompressed, not ROI-isolated). It is clear that compression as severe as 0.156 bpp (CF = 50, file size 2.0 KB, purple curve) still allows remarkably good iris recognition performance. For example, the FRR remains below 1% at an FAR of 1 in 100,000. It seems extraordinary that image arrays recovered from as little as 2,000 bytes of data are still so



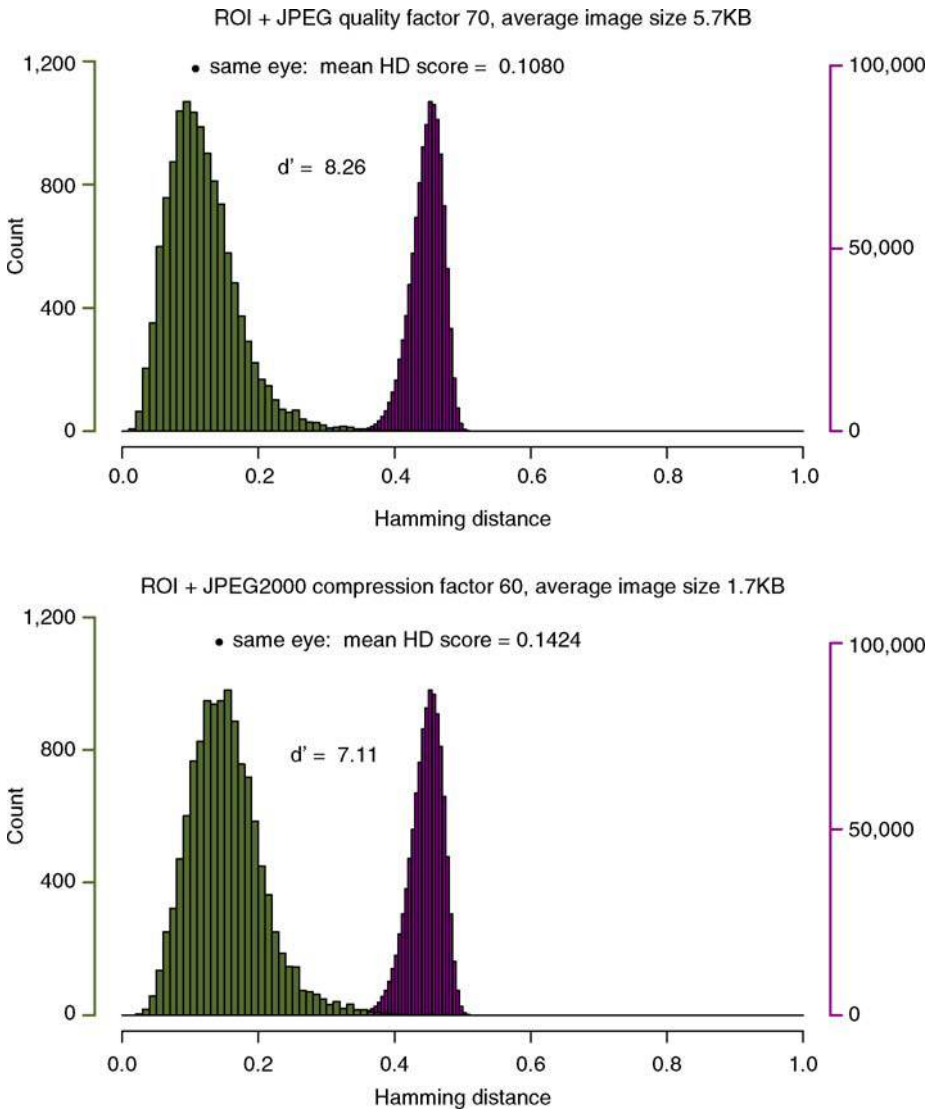
Iris Recognition Performance Under Extreme Image Compression. **Figure 4** ROC curves and data size statistics showing the consequences of ROI isolation prior to JPEG image compression, so that the available information budget is allocated almost entirely to the iris texture itself. The same quality factors were specified as in the corresponding curves of Fig. 3, and the recognition performance is generally comparable, but now the data reduction factors achieved in each case are twice as great.



Iris Recognition Performance Under Extreme Image Compression. **Figure 5** ROC curves and data size statistics showing iris recognition performance when the cropped and ROI-isolated images are compressed using JPEG2000 at various compression factors. Performance with file sizes of merely 2,000 bytes (CF = 50, *purple curve*) remains remarkably unimpaired compared to baseline (*black curve*); but further compression begins to exact a high toll (*blue-green curve*).

serviceable for iris recognition. It is possible that part of the explanation lies in the similarity between the Daubechies wavelets used for the DWT in JPEG2000 coding, and the Gabor wavelets used in the creation [10] of the IrisCode itself, so that information lost in such severe compression is not used in the IrisCode anyway. However, a watershed seems to exist at 2,000 bytes, since a pronounced degradation becomes evident when images are further compressed to 1,700 bytes (CF = 60, blue-green ROC curve in Fig. 5).

As with all biometrics, ROC curves as plotted in Figs. 3–5 reflect the overlapping tails of the two distributions of similarity scores computed for images from same or different eyes. For the present work, the similarity score is a normalized Hamming Distance (HD), which is the fraction of bits disagreeing between two IrisCodes among the bits compared [10]. It is also informative to see the full distributions of HD scores, which are presented in Fig. 6 for two of the compression schemes. In each panel, two different ordinate axis scales are used to facilitate visualization since there are 1,002,386 counts in the "all against all other" distribution (magenta) created by comparing different eyes, but only 12,214 counts in the distribution (olive) made by all same eye comparisons across the database. The upper panel shows the distributions obtained with ROI+JPEG compression at QF = 70, which created an average file size of 5,700 bytes. The recognition performance obtained with that compression scheme was almost indistinguishable from the baseline performance (black ROC curve: no compression, ROI, or cropping). The lower panel shows the distributions obtained with ROI+JPEG2000 compression at CF = 60, which created an average file size of just 1,700 bytes and generated the blue–green ROC curve in Fig. 5. It is remarkable that such extremes of compression do not have catastrophic effects on the separability of the pair of distributions. Instead, as seen in Fig. 6, the distribution obtained from different eyes (magenta) is virtually unchanged, whereas the



Iris Recognition Performance Under Extreme Image Compression. **Figure 6** Distributions of Hamming Distance scores comparing same and different eyes in the NIST database, for two of the image compression schemes bracketing the range of schemes studied. Even in the most severe case (*lower panel*) using images compressed to only 1,700 bytes, the dual distributions have little overlap and so decisions about identity remain robust.

distribution obtained from same eye images (olive) is shifted to the right by a small amount, corresponding to an increase in the mean HD score from 0.1080 to 0.1424 as indicated by the two dots and a projected vertical line for comparison.

Information theory provides certain metrics for defining the "distance" between two random variables in terms of their entire probability distributions. When both random variables are distributed over the same set of possible outcomes, such as the HD scores that were tallied in the histograms for same and for different eyes in Fig. 6, then the relative entropy or Kullback–Leibler distance is a natural way to measure the overall distance between the two distributions. As a measure of separation, it is also called the "information for discrimination." Unfortunately, this measure becomes undefined if there are some values that only one random variable can have while other values are accessible only to the other random variable. Since the distributions of HD scores obtained from comparisons

between different eyes in Fig. 6 vanish for scores smaller than about 0.3, and likewise the score distributions for same eyes attenuate to zero over much of the other distribution, the calculated Kullback–Leibler distance between these distributions is infinite and meaningless, unless based on nonvanishing theoretical models for them or by adding arbitrary quantities that then become decisive for this metric. An alternative family of distance metrics, encompassing the Fisher ratio and *Z*-scores, define distance in terms of the difference between the means of the two distributions, normalized by some function of their standard deviations. One such is the d' metric of decidability in signal detection theory, defined as $d' = |\mu_1 - \mu_2| / \sqrt{\frac{1}{2}(\sigma_1^2 + \sigma_2^2)}$, where μ_1 and μ_2 are the means and σ_1 and σ_2 are the standard deviations. A limitation of this metric is that by considering only the first two moments of the distributions, it makes no explicit use of skew, kurtosis, and higher moments that are more sensitive to mass in the tails. Thus d' might be said to take a "Gaussian view" of the world, whereas the skewed distributions in Fig. 6 are clearly not Gaussian. Nonetheless, the d' scores for each underlying pair of distributions obtained with each of the compression schemes studied here have been included within the ROC graphs in Figs. 3–5. They show a small but systematic trend of deterioration with more aggressive levels of image compression. But as is clear from the two bracketing extremes presented in Fig. 6, the separability of the two underlying distributions remains remarkable, despite the massive compression factor reaching 180:1 reduction from the original images.

Another metric system often used in decision theory to summarize overall performance by a single scalar statistic is the *area* under the ROC curve. Clearly a value of 1.0 represents perfection since it arises only from the complete absence of overlap between the two distributions. The ten different ROC curves plotted in Figs. 3–5 appear to have significant amounts of missing area, but this is an illusion due to the logarithmic abscissa and the magnified ordinate ranging just between 0.95 and 1.00. In fact the area under the baseline black ROC curve (present in all three figures) is 0.999985, and the areas under most of the other nine curves are reduced from this value in only the sixth decimal place.

Summary

From studying the effects of three schemes for image compression on iris recognition performance, one is drawn to the surprising conclusion that even images compressed as severely as 150:1 from their original full-size formats, to just 2,000 bytes, remain very serviceable. It is important to use region-of-interest isolation of the iris within the image so that the coding budget is allocated almost entirely to the iris texture rather than to costly irrelevant structures such as eyelashes; and it

Iris Recognition Performance Under Extreme Image Compression. Table 1. Summary of the compression schemes, resulting file sizes, and their effects on computed IrisCodes, expressed as entropy per code bit and as the fraction (HD) of bits that were changed from those computed for the original full-size images

Strategy	Compression parameter	Average image size	Entropy/code bit	Interoperability HD
Cropping (320×320) + JPEG compression	QF = 70	12.4 KB	0.053 bit	0.006
	QF = 30	5.7 KB	0.087 bit	0.011
	QF = 20	4.2 KB	0.147 bit	0.021
Cropping + ROI + JPEG compression	QF = 70	5.7 KB	0.112 bit	0.015
	QF = 30	2.7 KB	0.147 bit	0.021
	QF = 20	2.1 KB	0.199 bit	0.031
Cropping + ROI + JPEG2000 compression	CF = 20	5.1 KB	0.130 bit	0.018
	CF = 50	2.0 KB	0.179 bit	0.027
	CF = 60	1.7 KB	0.219 bit	0.035

is important to use JPEG2000 instead of JPEG as the compression protocol. Advantages of this overall approach from the perspective of Standards bodies and interoperability consortia are that the compact image data (when decompressed) are a native rectilinear arrays; no proprietary methods are required; and the distortions that can arise from alternative coordinate transformation methods such as polar unwrapping or polar sampling are avoided.

As concluding measures, the IrisCodes generated under each scheme were compared with those generated for the corresponding original uncompressed images. The entropy $H = -\sum_i p_i \log_2(p_i)$ or uncertainty per code bit caused by each compression scheme is tabulated in Table 1. For reference, the entropy associated with the states of bits in IrisCodes calculated from different images of the same eye, due merely to variation in image capture, is typically 0.506 bit; Table 1 shows that the corrupting effect of the image compression schemes is much less than this native uncertainty in the bits of IrisCodes for a given eye. The final column of Table 1 tabulates, as interoperability scores, the average HD (fraction of disagreeing bits) between the IrisCodes obtained before and after image compression for each scheme and for each compression parameter. They indicate that only about 2–3% of the IrisCode bits change as a consequence of image compression even as severe as to 2,000 bytes. When considered in the context of Fig. 6 showing the HD distributions for same and different eyes, it is clear that an increment of 0.02–0.03 in HD score is a negligible impact indeed. In conclusion, it appears that rough convergence between data length and standard description length for this biometric system is possible. These observations appear to vindicate the applicability to biometrics of the fundamental insights of Shannon [3] and Kolmogorov [4] and the relevance of their analyses of asymptotic compressibility.

Related Entries

► Iris Encoding and Recognition Using Gabor Wavelets

References

1. Bradley, J., Brislawn, C., Hopper, T.: The FBI Wavelet/Scalar Quantization standard for grayscale fingerprint image compression. Proc. SPIE (Applications of Digital Image Processing XIX) **2847** (1996)
2. Daugman, J., Downing, C.: Effect of severe image compression on iris recognition performance. IEEE Trans. Inform. Forensics Secur. **3**, 52–61 (2008)
3. Shannon, C.: A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423 (1948)
4. Kolmogorov, A.: Three approaches to the quantitative definition of information. Probl. Inform. Transm. **1**, 4–7 (1965)
5. Terzopoulos, D., Waters, K.: Analysis and synthesis of facial image sequences using physical and anatomical models. IEEE Trans. Pattern Anal. Mach. Intell **15**, 569–579 (1993)
6. Cappelli, R., Maio, D., Maltoni, D.: Synthetic fingerprint-image generation. Proc. Int. Conf. Pattern Recognit. **15**, 475–478 (2000)
7. Cui, J., Wang, Y., Huang, J., Tan, T., Sun, Z.: An iris image synthesis method based on PCA and super-resolution. Proc. 17th Int. Conf. Pattern Recognit. **4**, 471–474 (2004)
8. Zuo, J., Schmid, N., Chen, X.: On generation and analysis of synthetic iris images. IEEE Trans. Inform. Forensics Secur. **2**, 77–90 (2007)
9. Daugman, J., Downing, C.: Epigenetic randomness, complexity, and singularity of human iris patterns. Proc. R. Soc. B Biol. Sci. **268**, 1737–1740 (2001)
10. Daugman, J.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**, 21–30 (2004)
11. Rakshit, S., Monro, D.: An evaluation of image sampling and compression for human iris recognition. IEEE Trans. Inform. Forensics Secur. **2**, 605–612 (2007)
12. National Institute of Standards and Technology. Iris challenge evaluation. http://iris.nist.gov/ice/
13. Wallace, G.: The JPEG still picture compression standard. Commun. ACM **34**, 30–44 (1991)
14. International Organisation for Standards: Information technology – Digital compression and coding of continuous-tone still images. ISO/IEC **10918** (1994)
15. Bradley, A., Stentiford, F.: JPEG2000 and region of interest coding. In: Digital Imaging Computing Techniques and Applications. Melbourne, Australia (2002)
16. International Organisation for Standards: Information technology – JPEG2000 image coding system. ISO/IEC **15444-1** (2004)
17. Christopoulos, C., Skodras, A., Ebrahimi, T.: The JPEG2000 still image coding system: an overview. IEEE Trans. Consum. Electron **46**, 1103–1127 (2000)
18. Hsu, R., Griffin, P.: JPEG region of interest compression for face recognition. IDENTIX Doc. **RDNJ-04-0102** (2005)
19. Registered Traveler Interoperability Consortium (RTIC): Technical Interoperability Specification for the Registered Traveler Program. http:///www.rtconsortium.org/

Iris Recognition Systems

► Iris Acquisition Device

Iris Recognition Using Correlation Filters

Yung-hui Li[1], Marios Savvides[2], Jason Thornton[2], B. V. K. Vijaya Kumar[2]
[1]Language Technology Institute, Carnegie Mellon University, Pittsburgh, PA, USA
[2]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Synonym

Iris Recognition with deformation and occlusion estimation

Definition

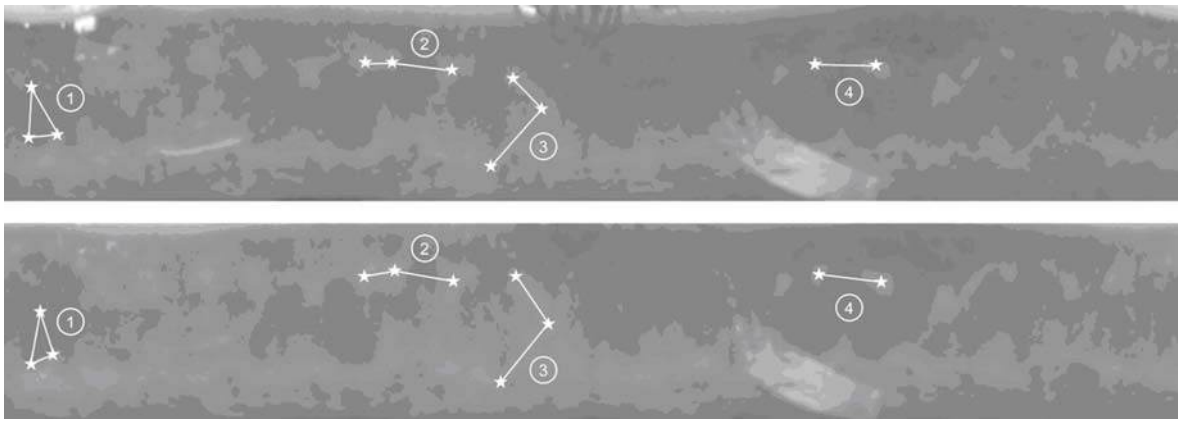
Algorithms for iris recognition usually consist of applying feature extraction on raw iris pattern, then matching against features. However, two important techniques in machine learning and pattern recognition, namely probabilistic graphical model, and advanced correlation filters, have not been used for iris recognition. By using probabilistic graphical models for iris texture deformation, with the observation being the correlation output derived from applying correlation filters to local iris regions, problems of iris pattern local deformations and occlusions can be handled and recognition performance can be improved over that of the conventional iris recognition algorithms.

Introduction

In the past two decades, iris recognition has emerged as one of the most promising modalities for biometric recognition. Many algorithms have been proposed to improve the recognition performance of iris recognition. However, there are still some obstacles which keep us from achieving near-perfect recognition results. Those obstacles include:

1. Local deformation in iris texture: iris texture may deform locally because of the dilation, contraction, or motion of the pupil, as shown in Fig. 1. Note that these deformations appear locally, not globally. Therefore, a simple image global transformation can not solve this problem. A more advanced modeling technique is needed.
2. Occlusion: very often, iris texture is occluded by many different objects, for example, eyelid, eyelashes, or eye glasses. Proper estimation of occluded region is important for achieving high recognition rate iris recognition.

The problem of deformation and occlusion estimation is seldom addressed in iris recognition literature. In this article, this problem is addressed by two advanced techniques in the field of pattern recognition and



Iris Recognition Using Correlation Filters. **Figure 1** Example images to show local deformation of iris texture. (**a**) original iris texture (**b**) iris texture after local deformation. There are four regions where local deformation is observed. Every anchor point is marked with a white star, and local anchor points are connected with white lines. Note that the deformations are irregular with different direction in different region. Therefore, it cannot be solved by a simple image global transformation.

machine learning: probabilistic graphical models and advanced correlation filters. Interested readers are referred to the original publication for more details [1–3].

Data Preprocessing

Iris images have to be preprocessed before going into the framework of deformation and occlusion estimation. The stage of data preprocessing includes iris segmentation and coordinate transformation, as described in the entry "Iris Overview." After preprocessing, the iris texture map is transformed to the polar coordinate plane, as shown in Fig. 2.

Deformation and Occlusion Estimation

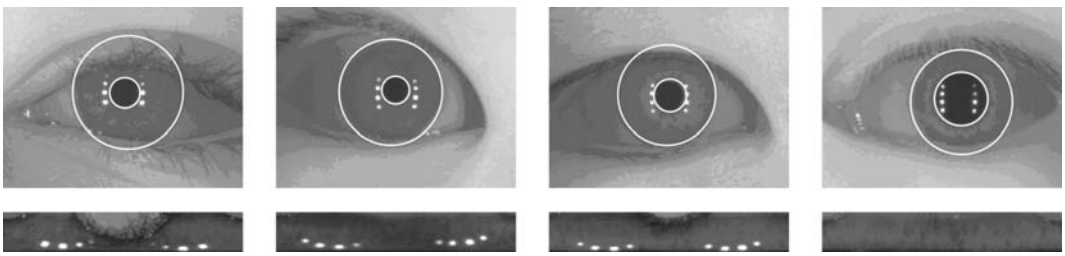
We can model the deformation and occlusion in iris pattern with a probabilistic graphical grid, as shown in Fig. 3. In Fig. 3a, an iris texture image is shown. The iris texture is dissected into smaller patches in order to observe the local deformation in more detail. In Fig. 3b, another iris texture image from the same class, is compared with Fig. 3a. The comparison is done in a patch-wise fashion. Every local patch is compared against the patch in the same location in the other image. The vectors in the center of each patch show how much and in what direction the patch is deformed. In Fig. 3b, we can see that the distance and the direction of the deformation vector is distributed randomly, and there is no global consistency among them.

In order to model such local deformation and occlusion, a probabilistic graphical model, arranged in a grid structure, is proposed, as shown in Fig. 3c. Suppose the number of patches in an iris image is N_s. Each node $\mathbf{d}_i$ in the model, $i = 1, \ldots, N_s$, represents a 2-D discrete-valued shift vector, the true value of which is hidden. The components of $\mathbf{d}_i$ are the vertical and horizontal shifts (in pixels) of the template region relative to the corresponding query region. The nodes ω_i are hidden binary-valued occlusion variables, where =0 and =1 denote that the region is occluded and un-occluded by the eyelid, respectively. Nodes O_i represent the observations, which include the match score of each local patch, and the occlusion statistics π_i.

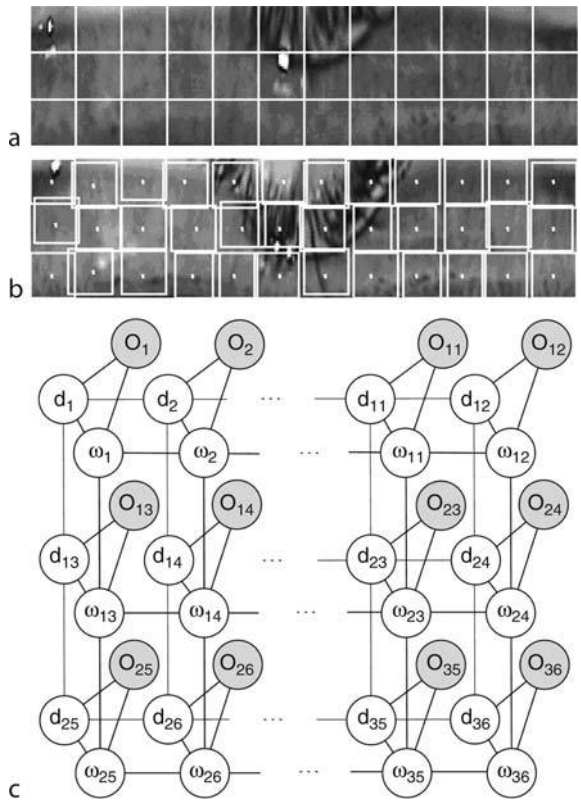
Match Score Computed from Correlation Filters

Computation of the match score has to be robust to the noise. At the same time, since we are estimating local deformation (relative shift) of each patch, it would be even better if the match score itself incorporates the information about the relative in-plane translation of the compared pattern. Advanced correlation filter is one of the techniques in pattern recognition field that can meet such requirement.

Theories of advanced correlation filters have been developed in the last two decades and they have been successfully applied in the field of biometrics [4–6]. There are many different versions of correlation filters and each of them is designed according to different criteria [7–11]. The one which is introduced in this article is the Optimal Trade-off Synthetic Discriminant Function (OTSDF), one of the most robust correlation filters [12]. The process of creating and using OTSDF is illustrated in Fig. 4. During training stage (i.e., creation of OTSDF), a few training images are given. From those training images,



Iris Recognition Using Correlation Filters. **Figure 2** The data preprocessing stage. First row: the raw eye images with iris boundary. Second row: the iris texture map after segmentation and coordinate transformation.



Iris Recognition Using Correlation Filters. Figure 3 Probabilistic graphical model for deformation and occlusion estimation. (**a**) the iris pattern dissected into patches (**b**) another iris pattern of the same class, dissected into patches. The vectors in the centers of the patches indicate the relative deformation of that patch. (**c**) the graphical model used for deformation and occlusion estimation.

OTSDF can be computed using a closed form expression in 2D Fourier domain. The expression is designed to minimize both the ▶ average correlation energy (ACE) and the ▶ output noise variance (ONV), but at the same time, make the output of the peak to be the preset value if the test image is one of the training images itself. Let $\mathbf{h}$ represent a vectorized form (e.g., by lexicographic scanning of the array into a column vector) of the correlation filter H(u, v), $\mathbf{x}_i$ represent the vectorized form of the ▶ Fourier transform of training image i, $\mathbf{D}$ represent a diagonal matrix with the average squared magnitude of all $\mathbf{x}_i$ along its diagonal, and $\mathbf{N}$ is the power spectral density of the input noise lexicographically re-ordered along its diagonal. The ACE and ONV can be represented as

$$\begin{aligned} ACE &= \mathbf{h}^{+}\mathbf{D}\mathbf{h} \\ ONV &= \mathbf{h}^{+}\mathbf{N}\mathbf{h} \end{aligned} \tag{1}$$

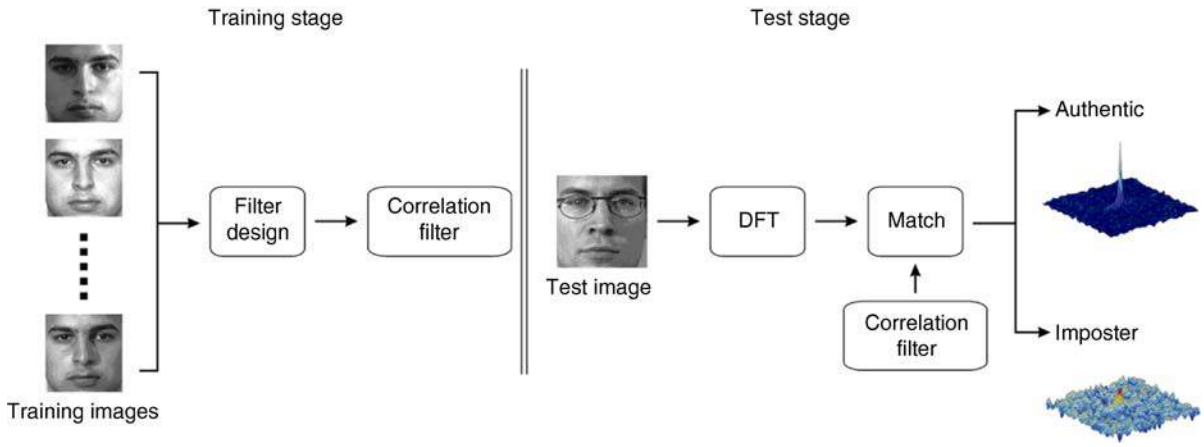
where + denotes complex conjugate transpose.

As stated above, the criteria of the filter design is to minimize both ACE and ONV subject to the linear peak constraints for the input training images. This is equivalent to minimizing a linear combination of the two terms (determined by a parameter α), subject to the same linear constraints. Lagrange optimization yields a closed from solution

$$\mathbf{h} = \mathbf{A}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{A}^{-1}\mathbf{X})^{-1}\mathbf{u} \tag{2}$$

where

$$\mathbf{A} = \alpha\mathbf{N} + \sqrt{1-\alpha^2}\mathbf{D} \tag{3}$$



Iris Recognition Using Correlation Filters. Figure 4 Left: The training stage of OTSDF; Right: the testing stage of OTSDF.

Matrix $\mathbf{X}$ contains $\{\mathbf{x}_i\}$ in its columns, and vector $\mathbf{u}$ contains the peak constraints (1 for authentics, 0 for impostors).

The OTSDF filter accomplishes two goals: it produces sharp peaks for training images, by reducing ACE, and it achieves tolerance to additive white noise, by reducing ONV in Eq. (1). The resulting filter gives good discrimination between authentic and impostor patterns, even when the authentic patterns are noisy.

Occlusion Metric

We need to have a metric for each local patch to evaluate the probability of this local patch belonging to an occluded region or a true iris pattern. We call this metric as "occlusion metric". For every local pixel x in a normalized iris map, we denote $\pi(x)$ as the occlusion metric.

There are some intuitive observations for estimating the occlusion metric for a particular patch. For example, if we observe the second row of Fig. 2 (the normalized iris map), we will see that in most cases, occlusion caused by eyelid happens in three regions: upper-left corner, upper-center region and the upper-right corner. Therefore, if the location of a point is close to these three regions, first, it is more likely to be an occluded region. Second, an occluded region caused by eyelids usually has higher value of pixel intensity. This is because in most cases eyelids look brighter than iris texture. Third, if we observe the occluded region in Fig. 3a and b, we would find that usually the eyelid region contains textures of irregular eyelashes. The line style of eyelashes in Cartesian coordinate is irregular. Some are straight, others are curved. After unwrapping them into polar coordinate, the shape of eyelashes become even more irregular. Therefore, they create irregular patterns on top of the eyelids. One special thing about eyelashes is their intensity is usually lower than iris texture. Therefore, eyelashes on top of eyelids lead to a large standard deviation of pixel intensity value in local region. This, too, can be used as one feature to distinguish occlusion from iris texture.

Based on the three observations mentioned above, we compute four features for every pixel on a normalized iris map. They are: (1) the mean intensity value on a local patch, centered around that pixel; (2) the standard deviation of the intensity value in that local patch; (3) the percentage of pixels whose intensity is greater than one standard deviation above the mean of the entire iris plane; (4) the shortest Euclidean distance from that pixel to one of the three "easily-occluded-region". After computing these features, we train a Fisher linear discriminant analysis (FLDA) model to compute a single value to represent the probability of this pixel being occluded. This FLDA model is trained by using 30 iris texture maps whose occluded regions are manually labeled.

Probabilistic Graphical Model

We have described the proposed probabilistic graphical model in previous section. Now, let us think about how to model the potential function of the nodes and between those nodes in Fig. 3c. As described earlier, for each patch, we have variables $\mathbf{d}_i$ and ω_i, to model the local shift and the occlusion, respectively. We can combine the both into one variable, say $\mathbf{h}_i = [\mathbf{d}_i, \omega_i]^T$, and $\mathbf{h}_i$ is a vector in three-dimensional space. For any pair of neighboring vectors $\mathbf{h}_i$ and $\mathbf{h}_j$, we can model the possibility of observing them $\Psi_{i,j}(\mathbf{h}_i, \mathbf{h}_j)$ with the following equation:

$$\Psi_{i,j}(\mathbf{h}_i, \mathbf{h}_j) = \Psi_{\mathbf{d},i,j}(\mathbf{d}_i, \mathbf{d}_j) \cdot \Psi_{\omega,i,j}(\omega_i, \omega_j) \quad (4)$$

The reason that we can decompose $\Psi_{i,j}(\mathbf{h}_i, \mathbf{h}_j)$ is that we assume the local shift of the patch and occlusion are statistically independent events. We can further model $\Psi_{\mathbf{d},i,j}(\mathbf{d}_i, \mathbf{d}_j)$ and $\Psi_{\omega,i,j}(\omega_i, \omega_j)$ with the following equation:

$$\Psi_{\mathbf{d},i,j}(\mathbf{d}_i, \mathbf{d}_j) = e^{\left\{-\frac{1}{2}\left(a\|\mathbf{d}_i\|+a\|\mathbf{d}_j\|+b\|\mathbf{d}_i-\mathbf{d}_j\|\right)\right\}} \quad (5)$$

$$\Psi_{\omega,i,j}(\omega_i, \omega_j) = \begin{cases} \alpha_0, \omega_i = \omega_j = 0 \\ \alpha_1, \omega_i = \omega_j = 1 \\ \alpha_2, \omega_i \neq \omega_j \end{cases} \quad (6)$$

Basically, Eq. (5) is assuming that the length of local shift vector should be small, the larger the length of the shift vector is, the smaller probability that it has. Besides this, the neighboring patches should have similar (if not exactly the same) local shift vector, so the probability is inversely proportional to the difference between the two local shift vectors. Equation (6) is to estimate the joint probability of the neighboring patches to be both occluded, both un-occluded, or one is occluded but not the other. These parameters can be trained by using a small portion of any iris database.

The next quantity we have to model is the potential function $\Psi_i(\mathbf{h}_i, \mathbf{O}_i)$ which is the probability of observing $\mathbf{O}_i$ when status of the node is $\mathbf{h}_i$. Obviously, $\Psi_i(\mathbf{h}_i, \mathbf{O}_i)$ must depend on ω_i because if a region is occluded then definitely it would affect the observation $\mathbf{O}_i$. Let us define the distribution $F(s)$ as the probability of having observed at least the true match score s, and $F(\pi)$ as the probability of having observed at least the true occlusion metrics π. Then $\Psi_i(\mathbf{h}_i, \mathbf{O}_i)$ can be expressed as:

$$\Psi_i(\mathbf{h}_i, \mathbf{O}_i) = \begin{cases} F_s(m(\mathbf{d}_i)), \omega_i = 0 \\ F_\pi(\pi_i), \omega_i = 1 \end{cases} \tag{7}$$

Where $m(\mathbf{d}_i)$denotes the match score when the given shift vector is $\mathbf{d}_i$. Since $F(s)$ is the probability of having observed at least the true match score s, it can be modeled as the cumulative distribution function (cdfs) of s and same thing applies to $F(\pi)$. Therefore, $F(s)$ and $F(\pi)$ can be expressed as:

$$F_s(S) = P(s < S) = \int_{-\infty}^{S} N(s; \mu_s, \sigma_s^2) ds \tag{8}$$

$$F_\pi(\Pi) = P(\pi < \Pi) = \int_{-\infty}^{\Pi} N(\pi; \mu_\pi, \sigma_\pi^2) d\pi \tag{9}$$

Where s represents the match score observed at the unknown true shifts for given un-occluded iris region, and π represents the occlusion metrics observed at the unknown true shift for given occluded iris region. Usually the distribution $P(s)$ and $P(\pi)$ are assumed to be normally distributed, as in Eqs. (8) and (9).

Final Score Computation

The equations we mentioned in last section are all about estimating probability distribution for each sub-region (patch). In order to generate a single value score for comparing two different irises, we have to combine scores from all sub-regions. We define the final Score M as:

$$M = \frac{\sum_{i=1}^{N_s} \beta_i M_i}{\sum_{i=1}^{N_s} \beta_i} \tag{10}$$

Where M_i is the score from each sub-region, and β_i is the weighting coefficient defined as the probability of sub-region i to be un-occluded, given the observation:

$$\beta_i = \hat{P}(\omega_i = 0 | \mathbf{O}) \tag{11}$$

The score of a single sub-region M_i is also computed based on a probabilistic point of view. Since we have estimated the joint probability distribution of the model parameter $\mathbf{h}_i$ and the observation $\mathbf{O}_i$, we should be able to compute the posterior probability distribution $\hat{P}(\mathbf{d}_i = \mathbf{d}|\mathbf{O})$. So M_i can be computed as follows:

$$M_i = \sum_{\mathbf{d}} m_i(\mathbf{d}) \hat{P}(\mathbf{d}_i = \mathbf{d}|\mathbf{O}) \tag{12}$$

Specifically speaking, the posterior probability $\hat{P}(\mathbf{d}_i = \mathbf{d}|\mathbf{O})$ can be estimated from $\hat{P}(\mathbf{d}_i|\mathbf{O})$, as shown in the following:

$$\hat{P}(\mathbf{d}_i|\mathbf{O}) = \sum_{\omega_j} \hat{P}(\mathbf{h}_i|\mathbf{O}) \tag{13}$$

If we view Fig. 3(c) as a Markov Random Field [13], with the potential function specified in previous section, we can use Loopy Belief Propagation [14] to estimate the conditional probability $\hat{P}(\mathbf{h}_i|\mathbf{O})$. Assume $\delta_{j\to k}^i(\mathbf{h}_k)$ is the message passing from node j to neighboring node k at the ith iteration. Then we can compute as shown below:

$$\begin{aligned} \delta_{j\to k}^i(\mathbf{h}_k) = &\sum_{\mathbf{h}_j} \Psi_j(\mathbf{h}_j, \mathbf{O}_j) \cdot \Psi_{j,k}(\mathbf{h}_j, \mathbf{h}_k) \\ &\times \prod_{l \in N(j)-k} \delta_{l\to j}^{i-1}(\mathbf{h}_j) \end{aligned} \tag{14}$$

Where $\Psi_j(\mathbf{h}_j, \mathbf{O}_j)$ and $\Psi_{j,k}(\mathbf{h}_j, \mathbf{h}_k)$ are defined in Eqs. (7) and (6), respectively, and $N(j)$ denotes the set of all neighboring nodes for j in the graphical model.

After i iterations, the conditional probability $\hat{P}(\mathbf{h}_j|\mathbf{O})$ can be computed as follows:

$$\hat{P}(\mathbf{h}_j|\mathbf{O}) = \frac{1}{z_j} \Psi_j(\mathbf{h}_j, \mathbf{O}_j) \prod_{k \in N(j)} \delta_{k\to j}^i(\mathbf{h}_j) \tag{15}$$

where the normalizing coefficient z_j is given by

$$z_j = \sum_{\mathbf{h}_j} \Psi_j(\mathbf{h}_j, \mathbf{O}_j) \prod_{k \in N(j)} \delta_{k\to j}^i(\mathbf{h}_j) \tag{16}$$

Result

The framework of iris deformation and estimation has been tested on ICE database [16]. For ICE database, there are two sub-sets. The first one, ICE Experiment 1 contains 1425 images of right irises from 124 users,

resulting in 12214 authentic comparison and 1002386 imposter comparisons. ICE Experiment 2 contains 1528 images of left irises from 120 users resulting in 14653 authentic comparison and 1151975 imposter comparisons.

The proposed algorithm is compared against a baseline algorithm, which does not contain the probabilistic framework for iris deformation. This model assumes that there is only a possible global shift in iris image. And the match score between two irises can be computed with the following equation:

$$m(\mathbf{x}) = \frac{1}{|s_t|}\sum_{\mathbf{y}\in s_t} c_t^T(\mathbf{y})c_q(\mathbf{y}-\mathbf{x}) \qquad (17)$$

where $\mathbf{x}$ is the global shift vector, $c_t(\mathbf{y})$ and $c_q(\mathbf{y})$ denotes the unshifted template and query iris codes, respectively, s_t is the support of the template iris code and $|s_t|$ is the total number of element in template. This method is similar to the method that Daugman proposed.

For each experiment, we computed the False Reject Rate (FRR) at different levels of False Accept Rate (FAR). Specifically, FRR is measured at FAR equals to 1%, 0.1% 0.01%, and 0.001%. The results are shown in Table 1.

From Table 1, we can see that the proposed framework of iris deformation and estimation has a much better performance than the baseline, which has the same power as the Daugman's algorithm. It shows that proposed framework is indeed effective and very useful in iris recognition problems.

Iris Recognition Using Correlation Filters. Table 1 False reject rates (FRRs) of the proposed algorithm and baseline algorithm in two iris database

Methods	CASIA			
	FAR = 1%	0.1%	0.01%	0.001%
Baseline	1.0%	2.3%	4.8%	9.6%
With deformation model	0%	0.1%	0.7%	1.9%
Methods	ICE Experiment 1			
	FAR = 1%	0.1%	0.01%	0.001%
Baseline	2.35%	3.45%	4.82%	6.26%
With deformation model	0.17%	0.33%	0.82%	1.56%
Methods	ICE Experiment 2			
	FAR = 1%	0.1%	0.01%	0.001%
Baseline	2.16%	3.02%	3.92%	5.25%
With deformation model	0.64%	0.94%	1.26%	1.91%

Summary

Iris pattern deformations and occlusions are two most prominent problems in the field of iris recognition. In the past, some methods have been proposed for the problem of occlusion detection, however, few have addressed the problem of local deformation of iris textures. In this article, we introduce one novel technique which combines the most advanced theories in machine learning and pattern recognition to solve this problem. By the framework of deformation and occlusion estimation, two iris images (from the same class) can be matched successfully even when one or both of them suffers the in-plane, local deformation. Experimental results have shown the power of the novel framework.

From the point of view of computational efficiency, the proposed framework only requires a few iterations of probability estimation, which consists of operations of simple addition and multiplication for a few variables. It does not take much longer time than the conventional iris recognition algorithm, which makes it suitable for iris recognition system running in real-time.

Related Entries

- ▶ Identification and Authentication
- ▶ Iris Encoding and Recognition
- ▶ Iris Recognition, Overview

References

1. Thornton, J.: Matching Deformed and Occluded Iris Patterns: A Probabilistic Model Based on Discriminative Cues. Dissertation, Dept. of Electrical and Computer Engineering, Carnegie Mellon University (2007)
2. Thornton, J., Savvides, M., Vijaya Kumar, B.V.K.: Enhanced iris matching using estimation of in-plane nonlinear deformations. In: Proceedings of SPIE Defense and Security Symposium, Vol. 6202, 62020E (2006)
3. Kerekes, R., Narayanaswamy, B., Thornton, J., Savvides, M., Vijaya Kumar, B.V.K.: Graphical Model Approach to Iris Matching Under Deformation and Occlusion. In: IEEE Conference on Computer Vision and Pattern Recognition, 2007 (CVPR '07). vol. June 2007, pp. 1–6, 17–22

4. Vijaya Kumar, B.V.K., Savvides, M., Venkataramani, K., Xie, C.: Spatial frequency domain image processing for biometric recognition. In: Proceedings of International Conference on Image Processing, September 2002, pp. 53–56
5. Savvides, M., Vijaya Kumar, B.V.K.: Efficient design of advanced correlation filters for robust distortion-tolerant face recognition. In: Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance, July 2003, pp. 45–52
6. Thornton, J., Savvides, M., Vijaya Kumar, B.V.K.: Robust iris recognition using advanced correlation techniques. In: Proceedings of International Conference on Image Analysis and Recognition, September 2005, pp. 1098–1105
7. Vijaya Kumar, B.V.K., Mahalanobis A., Juday, R.D.: Correlation Pattern Recognition. Cambridge University Press, Cambridge, UK (2005)
8. Vijaya Kumar, B.V.K.: Tutorial survey of composite filter designs for optical correlators. Appl. Opt. **31**, 4773–4801 (1992)
9. Vijaya Kumar, B.V.K., Mahalanobis, A., Takessian, A.: Optimal tradeoff circular harmonic function corr1elation filter methods providing controlled in-plane rotation response. IEEE Trans. Image Process. **9**(6), 1025–1034 (2000)
10. Kerekes, R., Vijaya Kumar, B.V.K.: Correlation filters with controlled scale response. IEEE Trans. Image Process. **15**(7), 1794–1802 (2006)
11. Mahalanobis, A., Vijaya Kumar, B.V.K., Casasent, D.: Minimum average correlation energy filters. Appl. Opt. **26**, 3630–3633 (1987)
12. Vijaya Kumar, B.V.K., Carlson, D.W., Mahalanobis, A.: Optimal trade-off synthetic discriminant function filters for arbitrary devices. Opt. Lett. **19**, 1556–1558 (1994)
13. Perez, P.: Markov random fields and images. CWI Quarterly **11**(4), 413–437 (1998)
14. Felzenszwalb, P., Huttenlocher, D.: Markov random fields and images. Int. J. Comput. Vis. **70**(1), 41–54 (2006)
15. ICE iris database. National Institute of Standards and Technology. http://iris.nist.gov/ICE/, 2005. Accessed 1 March 27, 2009
16. Daugman, J.: How Iris Recognition Works. IEEE Trans. Circ. Syst. video Tech. **14**(1), 21–30 (2004)

Iris Recognition with Deformation and Occlusion Estimation

▶ Iris Recognition Using Correlation Filters

Iris Retina Biometric Fusion

▶ Simultaneous Capture of Iris and Retina for Recognition

Iris Sample Synthesis

Natalia A. Schmid
Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV, USA

Synonym

Synthetic iris images

Definition

Iris image synthesis is a process of creating images of an iris by means of statistical and stochastic models, computer graphic tools or through manipulating, modifying or transforming parts or complete images collected from real irises. Since the size of synthesized datasets can be made arbitrary large, these data are suggested to be used for the purpose of extensive testing of the performance and efficiency of newly designed iris recognition algorithms.

Introduction

Iris as a biometric has been known for a long time. However, only in the recent years it has gained a substantial attention of both the research community and governmental organizations resulting in the development of a large number of new iris encoding and processing algorithms. Most of the designed systems and algorithms are claimed to have exclusively high recognition performance. However, since there are no publicly available large-scale and even medium-size datasets, only very few newly designed algorithms have undergone extensive testing. There are several datasets of frontal view iris images presently available for public use. A brief summary of these databases is provided in Table 1.

With the lack of data, two major solutions to the problem of algorithm testing are possible: (1) physically collecting a large number of iris images or (2) synthetically generating a ▶ large scale dataset of iris images.

This article addresses the second approach. The following sections summarize a number of techniques to synthesize iris images and characterize methods to evaluate the performance of generated images.

Iris Sample Synthesis. **Table 1** Public iris databases

Database name	Database size	# of classes	# of Images per class	Color or gray scale
CASIA-I	756	108	7	gray
CASIA-III-device1	1,200	60	20	Gray
CASIA-III-device2	1,200	60	20	Gray
CASIA-III-Interval	2,655	396	NA	Gray
CASIA-III-Lamp	16,213	819	NA	Gray
CASIA-III-Twins	3,183	400	NA	Gray
WVU-off-angle	560	140	4	Gray
WVU	2453	359	NA	Gray
UBIRIS	1,877	241	NA	Color
UPOL	384	128	3	Color
BATH	1,000	50	20	Gray

The purpose of a synthesized dataset is to provide an option to compare efficiency, limitations, and capabilities of newly designed iris recognition algorithms through their testing on a large scale dataset of generated irises. Although adding synthetic data to the test set, or adding artificial noise to data for a scenario testing may introduce a ► performance bias (see recommendations by Mansfield and Wayman [5]), when used with caution, a large synthesized dataset may evaluate efficiency and point to drawbacks of testing algorithms.

Available Models

The first methodology for generating synthetic iris images has been proposed by Cui et al. [1], where a sequence of small patches from a set of iris images was collected and encoded by applying Principal Component Analysis (PCA) method. Principal components were further used to generate a number of low resolution iris images from the same iris class. The low resolution images were combined in a single high resolution iris image using a superresolution method. A small set of random parameters was used for the generation of images belonging to different iris classes.

Another method for the generation of synthetic iris images based on the application of Markov Random Field (MRF) has been recently developed by Makthal, Shah, and Ross [4, 6]. The generation technique relies on texture primitives to synthesize new iris images. Texture primitives are small portions of iris texture selected from a host iris image. The generation technique is deterministic in the sense that it regenerates texture primitives locally, but positions them at locations specified by a frequency of their occurrence map. The map specifies the probability of occurrence of a texture primitive at various locations in a generated image. Thus generated image globally exhibits a different structure compared to host images. A few images generated using this approach are shown in Fig. 1.

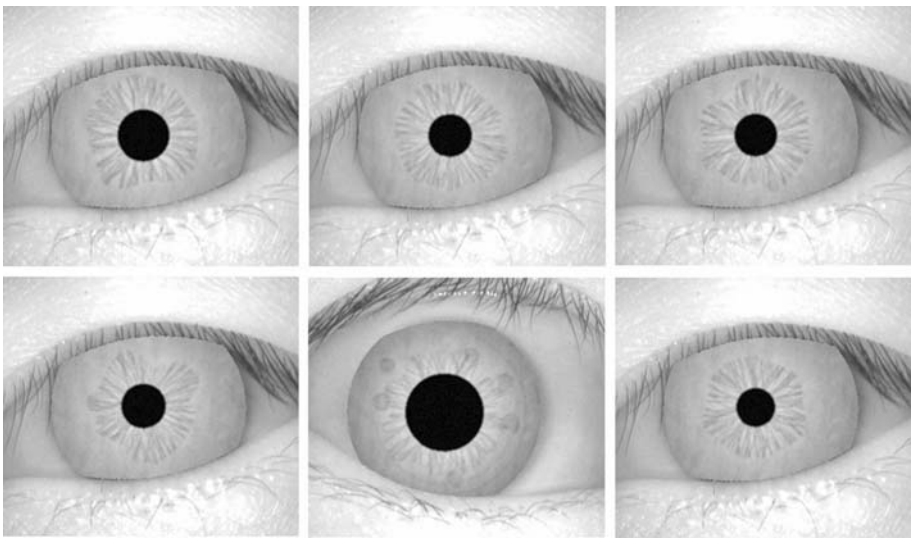
Lefohn et al. [3] developed an oculariste's approach using the computer vision technology for the purpose of both the ocular prosthetics and entertainment industries. In their work, a set of textured layers was used to render each iris.

Wecker et al. [7] combined characteristics of real irises to augment existing real iris databases. In their work a multiresolution technique known as reverse subdivision was used to capture the necessary characteristics.

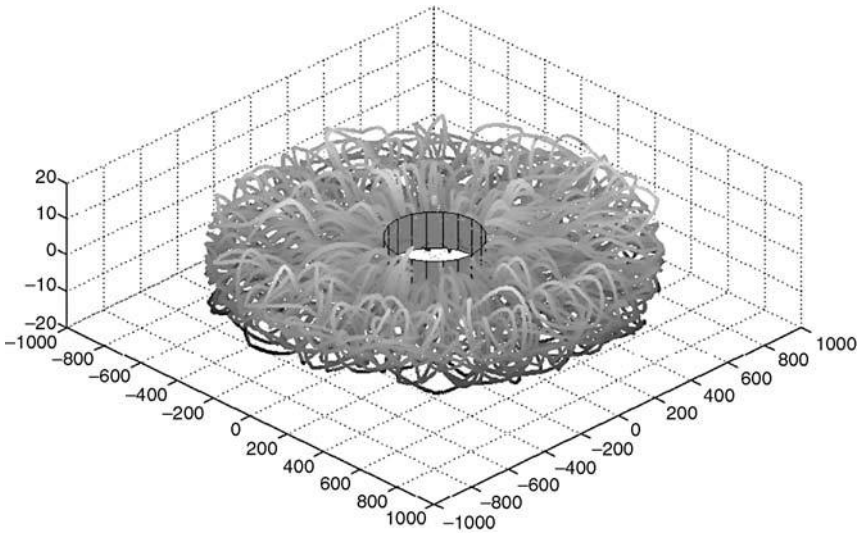
Zuo et al. [8, 2] took a model based, anatomy based approach for the generation of iris images. The work makes a number of observations on common visual characteristics of irises such as (1) radial fibers, radially arranged iris vessels, constitute the basis for iris tissues and dominate the structure information; (2) a large part of iris is covered by a semitransparent layer with a bumpy look and few furrows which are caused by retractor muscles; (3) the irregular edge of the top layer contributes to the iris pattern; (4) the collarette part is raised due to the overlap of sphincter and dilator muscles. Thus, the main frame of the iris pattern in this work is formed by the radial fibers, the raised collarette, and the partially covered semitransparent layer with the irregular edge. At the same time, the difference of pixel values in an infrared iris image is not only the result of the iris structure information. It is also related to the type of muscles, vessels, and cells that the iris is composed of, surface color, and lighting conditions. Involving all those visual and anatomical characteristics makes each synthetic iris look similar to a real iris.

To simulate the stochastic nature of individual iris patterns, the process of iris image generation is reduced to the generation of a dense fiber structure controlled by a set of random parameters and postprocessed using a variety of image processing techniques. The influence of controlled parameters is carefully researched and the operational range of parameters is evaluated.

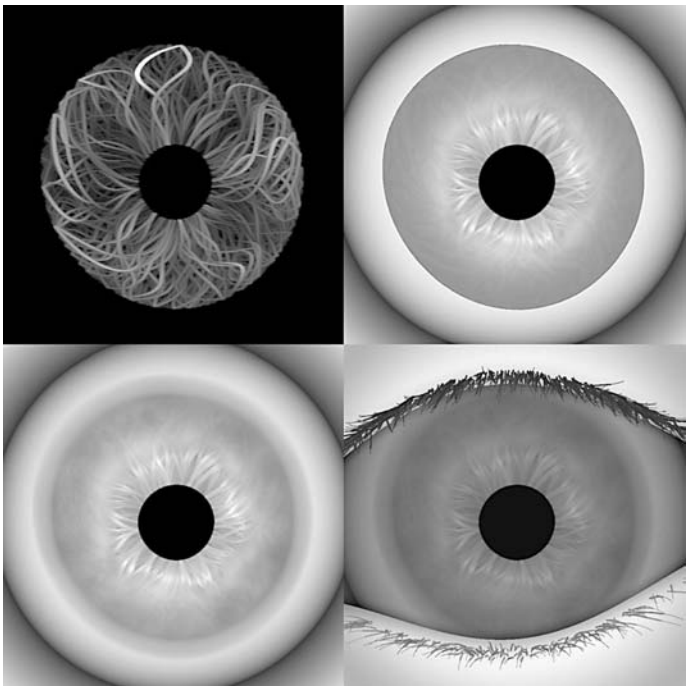
The value of each parameter is limited to a certain range to keep the common iris features. However, within the range it is allowed to vary maximally to increase the randomness of the iris pattern. A four step procedure for the generation of iris images was developed: (1) The generation of a continuous fiber structure in a cylindrical coordinate system; (2) Projection of the three-dimensional structure onto a two-dimensional image space based on the depth of fibers in the structure; (3) Modeling and blurring the top layer; and (4) Adding visual effects adherent to various parts of the iris. An example of a three-dimensional fiber structure is displayed in Fig. 2. The results of the procedures described in steps 2–4 are demonstrated in Fig. 3. A small gallery of generated iris images synthesized by following the approach by Zuo et al. is displayed in Fig. 4.



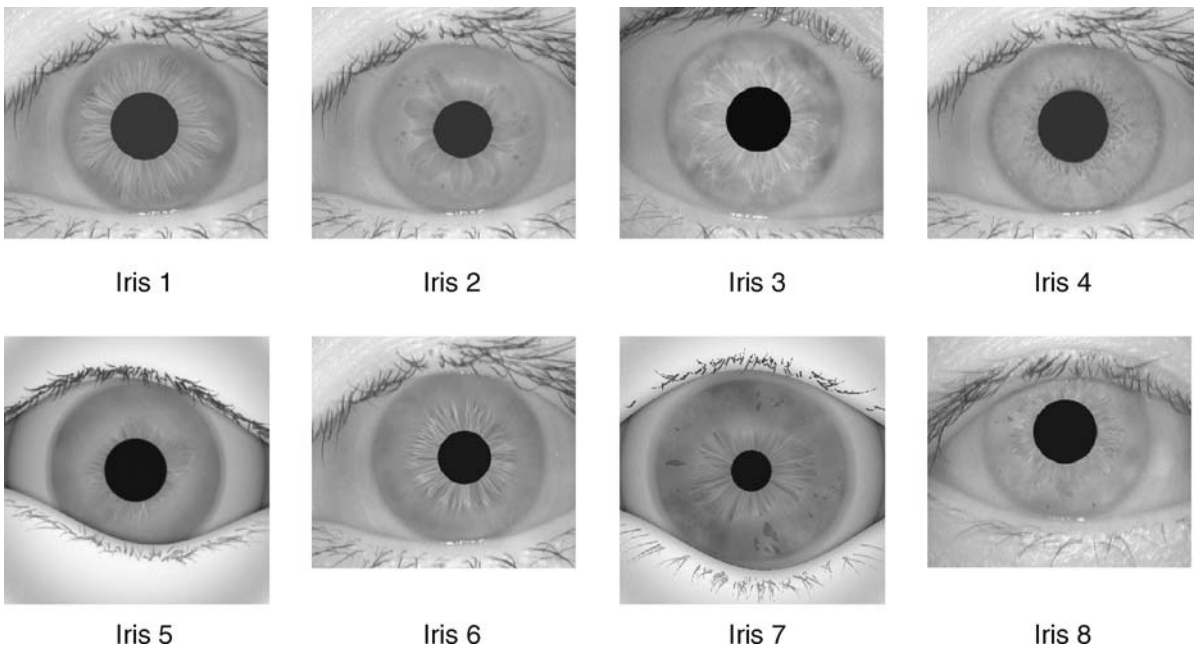
Iris Sample Synthesis. **Figure 1** A gallery of synthetic iris images generated using MRF approach. The regular structure imposed by MRF is smoothed by incorporating additional synthesized features such as furrows and crypts through a line integral convolution (Published with approval of S. Shah and A. Ross).



Iris Sample Synthesis. **Figure 2** An example of a three-dimensional fiber structure generated using the methodology described in Zuo et al. [2].



Iris Sample Synthesis. **Figure 3** Shown are the steps 2–4 of iris image generation.



Iris Sample Synthesis. **Figure 4** A gallery of synthetic iris images generated using model based, anatomy based approach. Iris 4 is a real iris image, a sample from CASIA dataset.

Quantifying Performance

When generating synthetic iris images, the problem that one faces is to define a measure of "realism." What is the set of requirements that a synthetic iris has to satisfy to be recognized and treated as a physically collected iris image? One can make various conclusions: (1) it should look like a real iris. (2) it should have

the statistical characteristics of a real iris. Real iris patterns are so anatomically complex that it is nearly impossible to mathematically describe any particular one. Thus, standards of realism will be limited to some degree.

The approaches that various research groups took in order to prove the validity of generated data varied substantially. For example, Makthal and Ross compared generated iris patterns against nonstochastic natural patterns from Brodatz library by invoking k-mean clustering technique. The data were clustered in two groups, iris and non-iris based on four distinct features associated with texture analysis and derived from the spatial grey level co-occurrence matrix, a second order statistics of images. Zuo et al. analyzed the verification performance of generated iris images by extrapolating the tails of genuine and imposter probability density functions fitted into genuine and imposter histograms of generated and real iris images. They also performed ► sensitivity analysis to conclude on importance of various parameters involved in generating iris images.

The common feature of all approaches to evaluating performance of generated iris data is that they focus on evaluating the similarity between real and synthetic iris images at three different levels: (1) global layout, (2) features of fine iris texture and (3) recognition performance.

Summary

Multiple approaches to generating sample iris images exist in the literature. These approaches vary in terms of tools and models used to synthesize images. Since synthetic data are known to introduce a bias that is impossible to predict, the data have to be used with caution. However, in the absence of a large- or even medium-scale dataset of real iris images, the generated data provide an option to compare efficiency, limitations, and capabilities of newly designed iris recognition algorithms through their testing on a large scale dataset of generated irises.

Related Entries

► Face Sample Synthesis
► Fingerprint Sample Synthesis

References

1. Cui, J., Wang, Y., Huang, J., Tan, T., Sun, Z.: An iris image synthesis method based on pca and super-resolution. In: Proceedings of 17th International Conference on Pattern Recognition (ICPR'04), Vol. 4, pp. 471–474. Cambridge, UK (2004)
2. Zuo, J., Schmid, N.A., Chen, X.: On generation and analysis of synthetic iris images. IEEE Trans. Inform. Forensics Security, **2**(1), 77–90 (2007)
3. Lefohn, A., Budge, B., Shirley, P., Caruso, R., Reinhard, E.: An ocularist's approach to human iris synthesis. IEEE Comput. Graph. Appl. **23**(6), 70–75 (2003)
4. Makthal, S., Ross, A.: Synthesis of iris images using markov random fields. In: Proceedings of 13th European Signal Processing Conference (EUSIPCO'05). Antalya, Turkey (2005)
5. Mansfield, A.J., Wayman, J.L.: Best practices in testing and reporting performance of biometric devices. http://www.cesg.gov.uk/site/ast/biometrics/media/BestPractice.pdf
6. Shah, S., Ross, A.: Generating synthetic irises by feature agglomeration. In: Proceedings of the International Conference Image Processing. Atlanta, GA (2006)
7. Wecker, L., Samavati, F., Gavrilova, M.: A reverse subdivision application. In: Proceedings of International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia (GRAPHITE'05), pp. 121–125. Dunedin, New Zealand (2005)
8. Zuo, J., Schmid, N.A.: A model based, anatomy based method for synthesizing iris images. In: International Conference on Biometrics (ICB'2006). pp. 428–435. Hong Kong, China (2006)

Iris Scan

► Biometric Sample Acquisition

Iris Scanner

► Iris Acquisition Device
► Iris Device

Iris Segmentation

Iris segmentation involves finding the pupillary and limbic boundaries of the iris within the image allowing

the iris region to be separated from the rest of the iris image. This step is necessary prior to the creation of an iris template and has significant impact on iris recognition performance.

► Iris Databases

Iris Segmentation Using Active Contours

Sung W. Park, Marios Savvides
Carnegie Mellon University, Pittsburgh, PA, USA

Synonym

Iris segmentation using snakes

Definition

Iris segmentation using active contours is finding an iris region in an image using *snakes* which are curves such as inner and outer boundaries at pupil and sclera. Iris segmentation is a key task for iris recognition. However, it is challenging to detect and exclude the true inner and outer boundaries of an iris. First of all, the occlusion caused by lower and upper eyelids and eyelashes makes it difficult to detect the iris region. Even though eyelids and eyelashes make no occlusion, the inner and outer boundaries are not exact circles, so circle detection such as the Hough transform is not enough to be applied. To obtain the iris boundaries, snakes have been successfully applied in literature. Snakes are energy-minimizing parametric closed curves guided by external forces.

Introduction

Among various biometric techniques, iris recognition can provide stable and accurate recognition rates, since irises have highly unique pattern consisting of complex tissue [1]. Also, iris pattern forms early, and remains the same throughout life, so iris recognition does not have aging or wearing problems unlike face or fingerprint recognition. However, for iris recognition, it is required to segment iris regions successfully, and moreover, iris segmentation is challenging in many reasons. First of all, the different lighting conditions often give shadows and specular reflections, and also, the radius of the pupil changes because of pupil dilation and contraction. Thus, we need an iris segmentation method robust on these changes in iris images. Moreover, unfortunately, we often have troubles to obtain a whole iris region because of the occlusion by eyelids and eyelashes. In particular, the upper outer boundaries are often partly or severely occluded by upper eyelids and eyelashes.

To get successful results of iris recognition in spite of these occlusion problems, the occlusion regions should be detected and excluded during segmentation. Active contours [2, 3] have been successfully applied to detect true iris regions by detecting the boundaries created by both an iris and occlusion.

Snakes for Curve Detection

In computer vision, edges or curves in an image are commonly used as features, since features give a strong presence in the images. For the low-level ► **feature detection**, the art of feature detection has been much studied, such as image filtering by digital convolution and simple thresholds. However, the results of these image filters dramatically change by thresholds, so more robust ways are required for reliable feature detection. To overcome these limitations of previous feature detectors, active shape modes with a variety of forms, principally *snakes* [3], have been proposed [4–7]. Snakes are energy-minimizing parametric closed curves guided by external forces. The aim of the snake is to find a location that minimizes energy for curve detection. Iris segmentation is dramatically enhanced by active contours, because active contours can fit noncircular boundaries in an iterative way. The basic concept of the snake algorithm is briefly summarized in this section.

Using snakes, an active contour is an ordered collection of points in an image:

$$P = \{p_1, p_2, \cdots, p_n\}, \quad (1)$$

where $p_i = (x_i, y_i)$ is a point in the contour. The snake method enables the points in the contour to approach

the boundary of an object iteratively through energy minimization. For each point in the neighborhood of p_i, an energy function is computed:

$$E = E_{int} + E_{ext}, \tag{2}$$

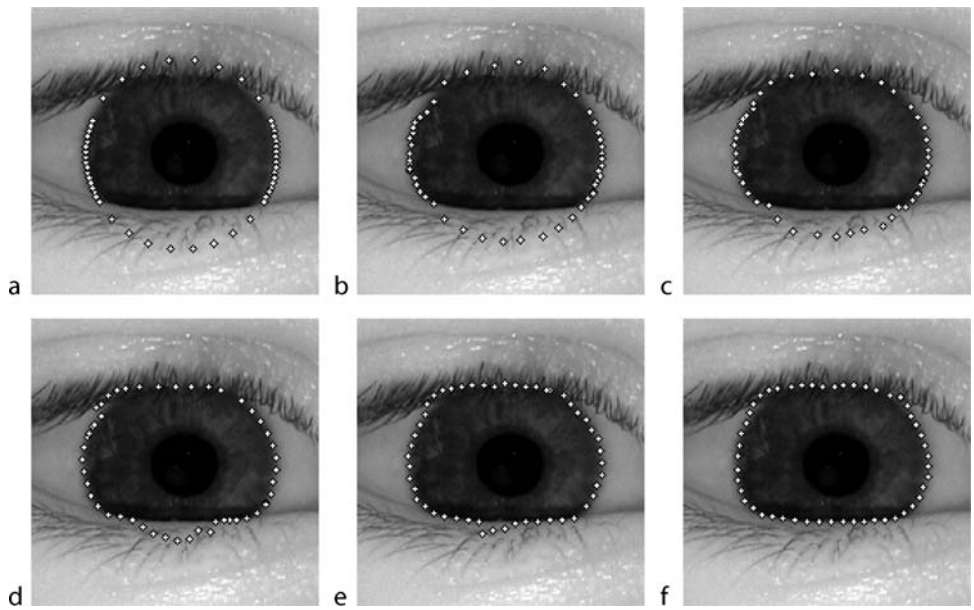
where E_{int} is the internal energy formed by the snake configuration, and E_{ext} is the external energy formed by external forces affecting the snake. First, E_{int} is dependent on the shape of the contour, and it is the sum of the contour continuity energy E_{cont} and the contour curvature energy E_{curv}. Next, E_{ext} is dependent on the image properties, such as the gradient, and is the sum of the image energy E_{img} and the energy of additional constraints E_{con}. α and β are constants providing the relative weighting of the two energy terms.

Generally, regardless of a variant of external constraint described in [3], the energy at every point can be written as

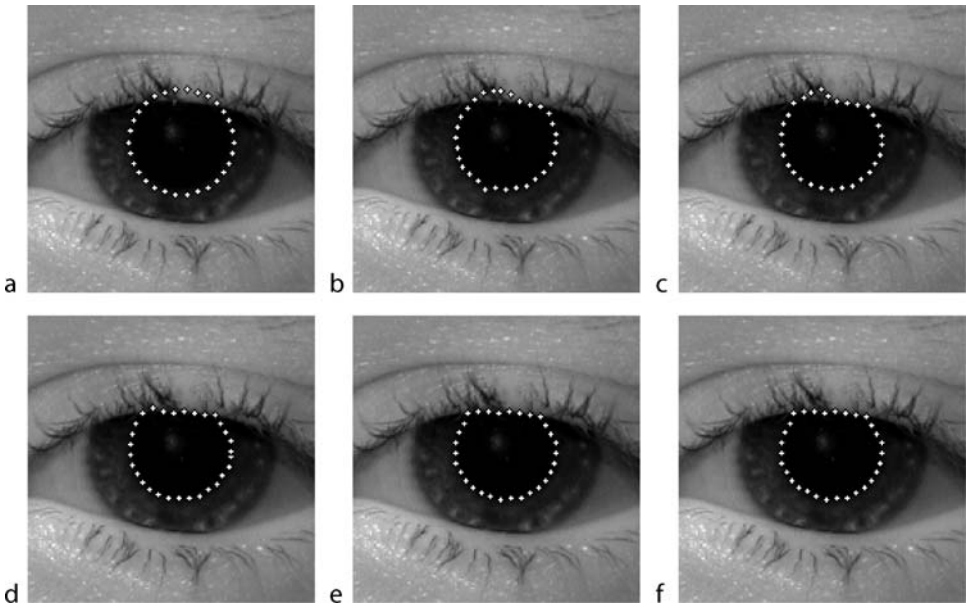
$$E_i = \alpha E_{cont,i} + \beta E_{curv,i} + \gamma E_{img,i} \tag{3}$$

where α, β, and γ are the weights of every kind of energy. The full snake energy is the sum of all the points. As α is bigger, snake points are more evenly spaced. Also, as β for a certain point increases, the angle between snake edges becomes more obtuse. Lastly, γ is responsible for making the snake point more sensitive to the image energy, rather than to continuity or curvature.

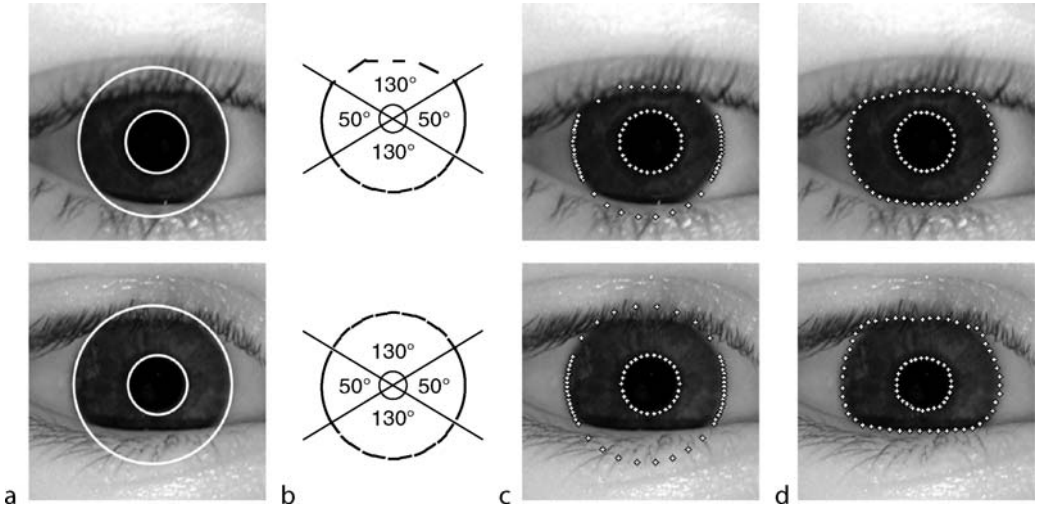
Figures 1 and 2 show two sets of points in the outer and inner boundaries. The initial points in Figs. 1a and 2a are set in the circles detected by a circle detector [7]. The snake algorithm updates the points iteratively and approach the boundaries. Commonly, a boundary is easier to be detected than an outer boundary, since the latter is often occluded more by eyelids or eyelashes than the former. So, the initial points in an outer boundary are selected in a more advanced way than those in an inner boundary. As shown in Fig. 2a, the initial points in an inner boundary are uniformly selected in the circle detected at pupil. On the other hand, the initial points in an outer boundary in Fig. 1(a) can be generated in either of the two ways shown in Fig. 3. The two rows in Fig. 3 provide two different templates for outer boundary detection. The first row in Fig. 3b shows the template when outstanding horizontal edges are detected around the upper region of the iris by the Sobel edge detector. In this case, we assume that an upper eyelid occludes the iris partially, and initializes the snake



Iris Segmentation Using Active Contours. **Figure 1** Snakes at *K*th iteration for detecting an outer boundary. (**a**) initial points; (**b**) $K = 1$; (**c**) $K = 2$; (**d**) $K = 5$; (**e**) $K = 12$; (**f**) $K = 30$.



Iris Segmentation Using Active Contours. **Figure 2** Snakes at *K*th iteration for detecting an inner boundary. (**a**) initial points; (**b**) $K = 1$; (**c**) $K = 2$; (d) $K = 5$; (**e**) $K = 12$; (**f**) $K = 30$.



Iris Segmentation Using Active Contours. **Figure 3** Segment an outer boundary at sclera. (**a**) the results of circle detection; (**b**) two templates of outer boundaries; (**c**) initial points generated by either of the two templates; (**d**) final results of snakes.

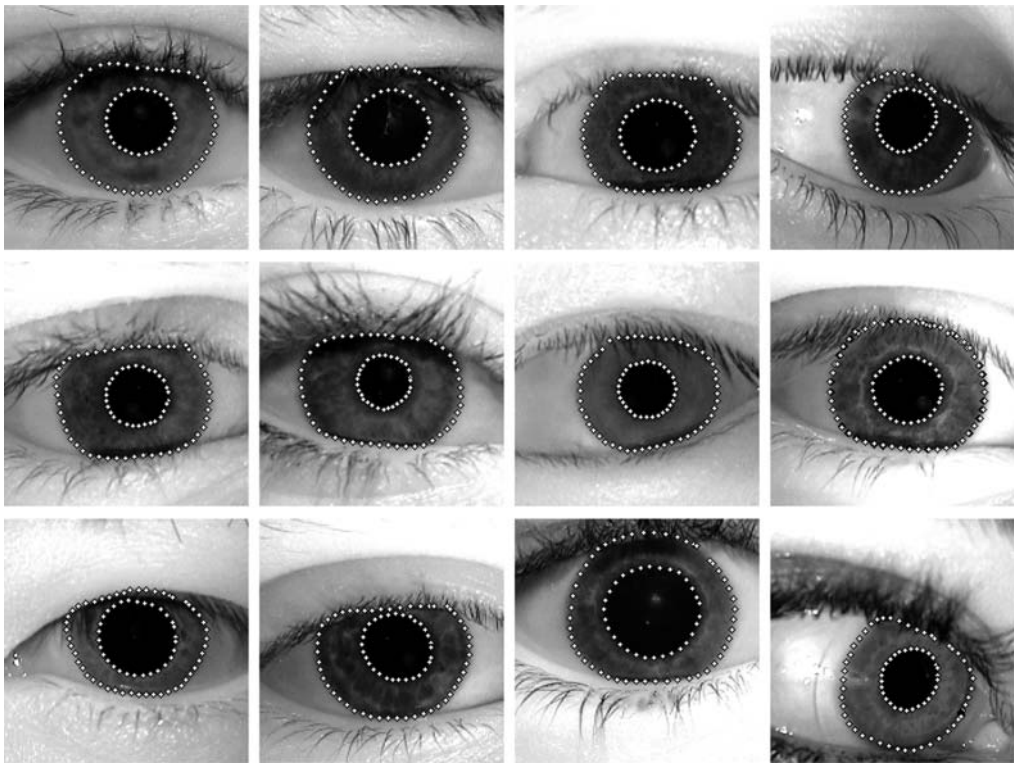
points in the lines connecting the detected edges. Otherwise, we let the initial points lie in the detected circle. In both cases, the snake points are initialized densely in the solid lines but sparsely in the dashed lines in the templates in Fig. 3; we assume that the outer boundaries around the solid lines, i.e., the left and right sides of the outer boundaries are more robust on the occlusion by eyelids. As shown in the final configuration of snakes in Fig. 3d, the snake points lie in the boundary almost uniformly in the end.

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix} \tag{4}$$

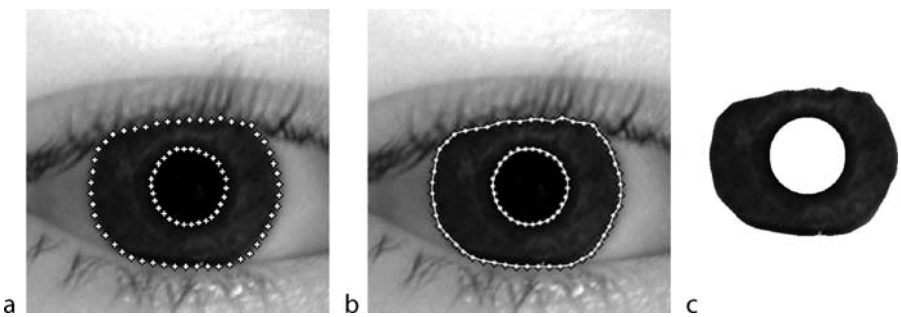
Segmentation Results

Figure 4 shows the results of iris segmentation proposed in this chapter. For experiments, the NIST ICE-1 database, which was constructed for NIST Iris Challenge Evaluation [8], is used. The segmentation results demonstrate that the proposed method using snakes produces reliable results in various cases. In particular, the segmentation results show that the proposed method is powerful to detect outer boundaries even when upper eyelids and eyelashes make severe occlusions inside iris regions. By active contours, the eyelids are excluded and the noncircular boundaries are detected successfully as shown in Fig. 4.

After the snake points converge, we can estimate the iris boundaries roughly. Next, we obtain more reliable boundaries by connecting the final snake points in an outer or inner boundary into a closed line, and smoothing the connected line as shown in Fig. 5.



Iris Segmentation Using Active Contours. **Figure 4** Examples of iris boundary detection using snakes.



Iris Segmentation Using Active Contours. **Figure 5** Segment a true iris region after applying snakes. (**a**) final results of snakes; (**b**) Smooth the boundary connecting the snake points; (**c**) final iris region.

Summary

For successful iris recognition, it is indispensable to segment the true iris region from an image. So, as a reliable iris recognition technique is demanded, the demand for robust iris segmentation also increases. However, iris segmentation is challenging in that upper and lower eyelids and eyelashes often make occlusion which is difficult to be detected by previous edge detectors or circle detectors. To improve the results of iris segmentation, active contours called snakes have been applied to iris boundary detection. In particular, snakes provide a powerful segmentation ability to detect and remove the occlusion by an upper eyelid which has been one of the most significant obstacles for reliable iris segmentation.

Related Entries

► Iris Recognition

References

1. Li, S.Z., Jain, A.K. (eds.): Handbook of Face Recognition. Springer, New York (2005)
2. Blake, A., Isard, M.: Active Contours. Springer, Heidelberg, Germany (1988)
3. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active Contour Models **1**, 321–331 (1988)
4. Bowyer, K., Hollingsworth, K., Flynn, P.: Image understanding for iris biometrics: A survey. University of Nortre Dame CSE Tech. rep., (2007)
5. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. 15(11) (1993)
6. Daugman, J.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 21–30 (2004)
7. Thornton, J., Savvides, M., Vijaya Kumar, B.: A bayesian approach to deformed pattern matching of iris images. IEEE Trans. Pattern Anal. Mach. Intell. **29**(1), 596–606 (2007)
8. Park, S.: National Institute of Standards and Technology. Iris Challenge Evaluation. http://iris.nist.gov/ice/

Iris Segmentation Using Snakes

► Iris Segmentation Using Active Contours

Iris Standards Evolution

► Iris Standards Progression

Iris Standards Progression

Dominique Harrington[1], Ryan Triplett[2]
[1]Tygart Technology Inc., Fairmont, MV, USA
[2]Biometric Services International, LLC, A National Biometric Security Project Company, Morgantown, WV, USA

Synonym

Iris standards evolution

Definition

Iris standards are used to provide a set of guidelines for iris system implementation, design, and interoperability. Both published and emerging standards for iris-based systems include recommendations for interoperability, data formats, conformance, compression, and quality.

There are two Technical Committees that create standards for the United States (U.S.) and international communities which solely concern biometrics. The American National Standards Institute (ANSI) International Committee for Information Technology Standards (INCITS) M1 Technical Committee creates U.S. national standards. This same group is the U.S. Technical Advisory Group for the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) Joint Technical Committee 1, Subcommittee 37 (SC37), which creates biometric international standards [1, 2]. Both ANSI/INCITS and ISO/IEC standards bodies ultimately publish the iris standards that are discussed within this chapter.

Introduction

Biometric standards development began in November 2001 with the creation of M1 in the United States.

Shortly thereafter, the international community began developing biometric standards with the creation of SC37 in June 2002. Each technical committee is broken into separate subcommittees. The overall structure of both organizations is similar with only the subcommittee names being different, as shown below:

- M1.1/WG1 – Task Group on Biometric Vocabulary
- M1.2/WG2 – Task Group on Biometric Technical Interfaces
- M1.3/WG3 – Task Group on Biometric Data Interchange Formats
- M1.4/WG4 – Task Group on Biometric Profiles
- M1.5/WG5 – Task Group on Biometric Performance Testing and Reporting
- M1.6/WG6 – Task Group on Cross Jurisdictional and Societal Issues

These subcommittees have created several standards governing the adoption of biometric technology throughout the industry. However, there are a few standards that specifically involve iris technology. These standards have either been published, or are emerging documents about to be published. The current versions are as follows:

- ANSI INCITS 379-2004 American National Standard for Information technology – Iris Interchange Format
- ANSI INCITS 1749-D: Part 6 -Conformance Testing Methodology for INCITS 379, Iris Image Interchange Format
- ISO/IEC 19794-6 Information Technology – Biometric Data Interchange Format – Part 6: Iris Image Data
- ISO/IEC 29109-6 (Base Document) Information Technology – Conformance Testing Methodology for Biometric Data Interchange Records as defined in ISO/IEC 19794 Biometric Data Interchange Format Standard – Part 6: Iris Image Data

The following sections will provide details concerning the specific standards that have been drafted and the evolution of standards development planned for the years to come.

Data Format Standards

Currently, data format standards are the only two standards published by both national and international bodies. They are the ANSI INCITS 379 Iris Image Interchange Format and the ISO/IEC 19794-6 Information Technology – Biometric Data Interchange Format – Part 6: Iris Image Data. Both of these standards are used to provide manufacturers an accepted data format to increase interoperability between multiple iris systems.

These standards identify two image interchange formats for biometric authentication systems that utilize the biometric iris modality. The first is ► rectilinear, consisting of a raw image, or a compressed image using the JPEG compression standard. The second is the ► polarized format, which requires pre-processing and image segmentation.

In addition to the interchange formats there are some specifications identified that are consistent throughout both standards. The following section outlines these specifications in detail.

Data Format Specifications

Each record obtained must have an iris record header which contains information about the image capture device and conditions by which the image was captured. The header will identify the number of features recorded, each iris as captured from the left or right, and the number of images captured from each eye. Each iris record should also contain an image sequence number with information about the quality and rotation of the image. An iris image data record will contain either rectilinear or polar format images and should not be mixed.

All images should have a 256 gray level range, allocating at least one byte per intensity value and provide at least 7 bits of useful intensity information. If ► specularity reflections from the illumination source occur, their intensity should be set to the saturation level or to a gray value of zero.

The eye should be illuminated by using near-infrared wavelengths of approximately 700 and 900 nm. The angle between a line extending from the center of the illumination source to the center of the pupil, and the optical axis of the iris camera should be at least 5° to prevent "red-eye" effect.

The iris image should have a minimum of 70 gray levels of separation between the iris and the sclera, and a minimum of 50 gray levels of separation between the iris and the pupil for any eye color. A minimum of 70% of the iris should be visible and not obscured by

specular reflections or any other obstructions. Table 1 outlines the quality recommendations by both the standards bodies while Table 2 discusses the definition of minimum requirement, medium quality, and high quality iris data records.

Conformance Standards

In accordance with the data format standards published, both M1 and SC37 are rapidly developing the conformance testing standards that will show compliance with the existing published standards. The standards are in their final development stages and are planned to be completed in 2008 [1, 2].

The conformance tests, in accordance with standards, produce verifiable results that a given manufacturer has designed a system capable of being interoperable with other iris systems available on the market. Such standards have also identified requirements for testing to maintain consistency.

Current Standard Bodies Developments

Like any emerging technology, new developments in the field of iris recognition are constantly being discovered, which have had direct effects on the standards community. In May 2007, the University of Cambridge released a technical report titled *Effect of severe image compression on iris recognition performance* [5]. This report documented three compression methods that retained rectilinear image formats compressing to as little as 2,000 bytes while still allowing quality recognition performance on publicly available iris image databases. The compression methods are as follows:

The first method reduced the size of the standardized iris image format by cropping the image and then compressed the cropped image into the JPEG format. "The standardized format is 640 by 480 pixels with 8 bits grayscale data per pixel, which uses approximately 307,200 bytes. Results show that a reduction

Iris Standards Progression. Table 1 Quality recommendations outlined in standards

Image Quality level	Image Quality value	Expected iris diameter, pixels	Minimum pixel resolution, pixels per mm	Optical resolution at 60% modulation, lp/mm	Comment
poor	0–25	–	–	–	Unacceptable quality
low	26–50	100–149	8,3	2,0	Marginal quality
medium	51–75	150–199	12,5	3,0	Acceptable quality
high	76–100	200 or more	16,7	4,0	Good quality

Iris Standards Progression. Table 2 Minimum, Medium, and High Quality Defined

Quality	Definition
Minimum Requirement	Image quality values between 0 and 25 and are used to indicate the image does not meet the minimum quality standards Images identified as being low quality use cameras with a minimum spatial resolution of 2.0 lp/mm at 60% or higher contrast, and pixel resolution at least 8.3 pixels per mm at the object plane [3, 4]
Medium Quality	Image quality values in the range of 51–75 are considered to be of medium quality Images identified as being medium quality use cameras with a spatial resolution of at least 3.0 lp/mm at 60% or higher contrast, and pixel resolution of at least 12.5 pixels per mm at the object plane [3, 4]
High Quality	Image quality values in the range of 76–100 are considered to be of high quality and should use cameras which have a spatial resolution of at least 4.0 lp/mm at 60% or higher contrast The pixel resolution should be at least 16.7 pixels per mm at the object plane [3, 4]

factor of 72:1 increased the Equal Error Rate (EER) by a factor of .0013."

The second method reduced the size of the image to an allocated Region of Interest (ROI). "This is done by substituting all non-iris parts of the image with a uniform grayscale; darker gray is used to signify eyelashes and lighter gray is used to signify the sclera. Highlighting the ROI at every Quality Factor has proven to be beneficial."

The third method reduced the data size using JPEG2000 compression. "This method is similar to the JPEG compression, but compresses the image further by a factor of 20–30%. JPEG2000 also allows use of a mask to specify an arbitrary shape (ROI) to control the allocation of the encoding budget. Results showed that a reduction factor of 180:1 using JPEG2000 compression and locating the ROI simultaneously increased the EER by .0016."

Publishing the technical report, moved SC37 WG3, at the Berlin meeting in June 2007, to propose the removal of the polar image format from the ISO/IEC standard because of defects caused by polar image distortion. M1 representatives were concerned about the impact the change would cause in the U.S. national biometric community. They recommended waiting until additional research and analysis was complete before making a final decision regarding removal of the polar format. The Department of Homeland Security (DHS) backed this decision by submitting a Position Paper in support of retaining the Iris Polar Image Data Format to M1 [6]. The paper outlined the effects on DHS's Registered Traveler (RT) program and general concerns with the technical report. As of December 3 2007 M1 representatives will submit their ballots, determining the acceptance or withdrawal of Project 1576-D – revision of INCITS 379–2004.

Future Standards

In addition to ANSI and ISO, there are other organizations that are concerned with biometric implementations both nationally and internationally. These organizations have created overall specifications and best practice documents that discuss biometrics as a whole. In some cases, these groups work with M1 and SC37 to establish future standards that meet the requirements of cross functional organizations outside the biometric industry mainstream. Table 3 identifies few of these organizations.

Summary

During the last six years, biometric standards have progressed rapidly compared with similar technology related industries from 6 in 2004 to the current 53 published standards. The current emerging and published

Iris Standards Progression. **Table 3** Other organizations affection standards bodies

Organization name	Brief overview
X9F4 (Financial)	The Accredited Standards Committee X9 (ASC X9) has the mission to develop, establish, maintain, and promote standards for the Financial Services Industry in order to facilitate delivery of financial services and products
International Labour Organization – ILO	The International Labour Organization (ILO) is the tripartite UN agency that brings together governments, employers, and workers of its member states in common action to promote decent work throughout the world
Institute of Electrical and Electronics Engineers – IEEE	The IEEE promotes the engineering process of creating, developing, integrating, sharing, and applying knowledge about electro and information technologies and sciences for the benefit of humanity and the profession
National Institute of Standards and Technology – NIST	Founded in 1901, NIST is a non-regulatory federal agency within the U.S. Commerce Department's Technology Administration. NIST's mission is to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life
National Physical Laboratory – NPL	The National Physical Laboratory (NPL) is the UK's National Measurement Institute and is a world-leading centre of excellence in developing and applying the most accurate measurement standards, science and technology

standards are proving to be only the beginning in the development of standards. Memberships in both groups are continuing to grow and future standards development is imminent.

Related Entries

- ▶ Biometrics System Design
- ▶ Interoperable Performance
- ▶ Iris Device
- ▶ Iris Encoding and Recognition
- ▶ Iris Image Quality
- ▶ Performance Evaluation, Overview

References

1. http://isotc.iso.org
2. http://m1.incits.org/
3. ANSI INCITS 379-2004 Iris Image Interchange Format.
4. ISO/IEC 19794-6:2005, Biometric Data Interchange Formats – Part 6: Iris Image Data.
5. "Effect of severe image compression on iris recognition performance", John Daugman, Cathryn Downing, May 2007, SC37 reference number 37N2125.
6. Position Paper in Support of Retaining the Iris Polar Image Data Format, 0733007 v0.1
7. M1 Letter Ballot, Document number M1/07-0633

Iris Super-Resolution

Yung-Hui Li[1], Marios Savvides[2]
[1]Language Technology Institute, Carnegie Mellon University, Pittsburgh, USA
[2]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

Synonyms

Iris image enhancement by Super-Resolution method; Super-Resolution for iris

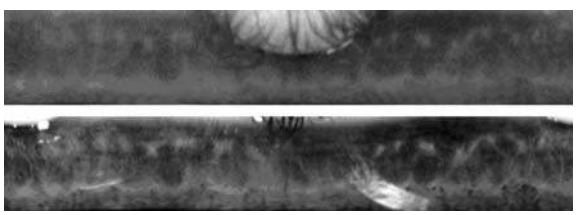
Definition

Super-Resolution is an image processing technique which takes input of a single or multiple low-resolution images and produces a single or multiple high-resolution images. By Super-Resolution processing, the quality of images can be enhanced and the follow-up stage of image processing (e.g., segmentation, object recognition, object tracking, or biometric identification) can achieve a higher success rate. The goal of iris Super-Resolution is to apply Super-Resolution technique in the specific domain as in iris image in order to enhance the quality of iris image. The iris image of better quality will result in a higher verification/recognition rate in iris recognition systems.

Introduction

Image resolution is a fundamental factor for the success of all kinds of image processing techniques, ranging from ▶ image segmentation, ▶ object recognition, tracking, 3D shape estimation, and reconstruction and biometric recognition. The higher resolution the input image has, the more accurate the output will be. Figure 1 shows the two iris images (in polar coordinate) which are taken from the same eye. The upper one is taken with a long distance camera. Because the distance between the eye and the camera is far, the resolution and the quality of the image is not very good. Therefore, many of the iris details cannot be shown in this picture. On the other hand, the lower image is taken with a high-resolution camera, placed closer to the eye. Comparing the image quality among these two pictures, one can clearly see that there are much more details in the second picture. When these two pictures are used as input to iris recognition system, the lower one will give a more reasonable score.

Although the maximal resolution of digital cameras has continually increased and the prices of them are



Iris Super-Resolution. Figure 1 Two iris images of the same eye, shown in polar coordinate. *Upper*: iris image taken from a long distance. Therefore, the image resolution and quality is not very high. *Lower*: iris image taken from a close camera, which contains much more details.

continually decreasing, there are still times that the region of object of interest in a picture is very small, and the resolution of the object is not high enough. In such circumstances, in order to enhance the object of interest and make the image of the object clearer, image processing technique needs to be applied to achieve this goal. Typically, in the field of image processing, those technique which takes input of one or multiple low-resolution images and output one or multiple high-resolution images are called image Super-Resolution.

Another problem that people have in their photos very often is the blurring of the images. The blur of the images might result from two reasons. One of them is the relative motion between the camera and the object. This is called motion blur. The other is the object is out of focus of the camera during image acquisition process. This is called focus blur. Besides image blur, another problem we need to overcome in order to have high quality image is the noise. Strictly speaking, the problems of image blur and image noises are different than the goal of Super-Resolution problem, however, in the literature, in order for Super-Resolution algorithm to be robust in real application, algorithm designers need to take into account the problem of image blur and image noises too.

Super-Resolution

Algorithms for Super-Resolution image can be categorized into two different types. The first is reconstruction-based algorithms [1–5]. In general, this type of Super-Resolution algorithm formulates the problem of Super-Resolution as a matrix inversion problem. Because digitized images consist of pixels, a digitalized image can be treated as a huge matrix, where each pixel associates with a real number value in a limited range. Therefore, a high-resolution image can be treated as a matrix of large size, while a low-resolution image can be treated as a matrix of small size.

A matrix can be converted to a vector, simply by concatenating all the columns (or rows) into a long vector. If we denote vector of the high-resolution image as X and the low-resolution image as Y, Super-Resolution problem can be simply formulated as

$$Y = MX + V$$

where M is a transformation matrix that represents the effect of imaging system and V is another column vector of the same size of Y, which represents the random noise inherent to any imaging system. M can be further decomposed into several factors to model different effects an imaging system produced to the original high-resolution image; those effects may include: relative motion between object and camera, camera blur effect, color filter effect, and down-sampling process.

The goal is to estimate the high-resolution image X from observed low-resolution image Y. This goal can be achieved by means of optimization. There are many different goals that can be chosen to be optimized; one of the most common ways is to minimize the ▶ L2 norm of the residual vector, which gives us the following equation:

$$\hat{X} = \arg\min_X \|Y - MX\|_2^2.$$

Different algorithms give different ways and different constraints to this optimization problem, and therefore, result in different solutions. In summary, reconstruction-based Super-Resolution algorithms formulate the problem as a matrix inversion problem and solve it by various optimization methods.

Second type of Super-Resolution algorithms is probability-based algorithms [6–10]. In this type of algorithm, the problem of reconstructing high-resolution image from low-resolution ones is modeled by a Bayesian approach. By Bayes' rule, the probability of Super-Resolution images given low-resolution ones Pr[Su| $\mathrm{Lo_i}$] can be decomposed into:

$$\Pr[Su|Lo_i] = \frac{\Pr[Lo_i|Su] \cdot \Pr[Su]}{\Pr[Lo_i]}.$$

Since $\Pr[Lo_i]$ is a constant and logarithm function is a monotonically increasing function, the goal of Super-Resolution can be achieved by this equation:

$$\arg\max_{Su} \Pr[Su|Lo_i] = \arg\max_{Su}(-\ln \Pr[Lo_i|Su] - \ln \Pr[Su])$$

Different algorithms are trying to learn $\Pr[Lo_i|Su]$ from training data with different methods, and integrated with different ways. But overall, the probability-based Super-Resolution algorithm is modeling the problem with a Bayesian approach and solves it with the ▶ maximum a posteriori (MAP) estimation.

Iris Super-Resolution

Iris Super-Resolution is a very new topic in field of image processing and biometric recognition. There is no relevant information found in the literature. In the following sections, a rough idea of how to do iris Super-Resolution is introduced to suggest a possible way of enhancing iris image quality by fusing information from multiple low-resolution images.

Information fusion can be thought of as a problem in creating a high-resolution iris image from multiple low-resolution images. Each one of the low-resolution images has part of information that is needed, and the goal is trying to extract those information needed and put them at the right position on two-dimensional space and finally create a pattern which gives an iris image of higher resolution.

The process of painting can be used as a metaphor to illustrate the process of iris Super-Resolution. When a very high-resolution painting is to be painted by first looking at a few low-resolution images and fill in more local details, the first thing is to select one best image from the given low-resolution image. Let it be called template image, and all other low-resolution image are scene images. Secondly, the template image is interpolated with zeros in order to enlarge to a bigger size. The new size of the is a parameter which can be fine tuned later. After the process of enlarging the image, zero-value pixel will spread evenly in entire image. They are called holes.

After the bigger image with holes is created, next step is trying to fill up every hole with appropriate numerical values, which are derived from combining useful information from other low-resolution images. In most cases, the appropriate numerical values for a hole can be inferred from the regions that surround it, and inside this region, the farther the points are away from this hole, the smaller impact they have toward this hole. Therefore, it is better to use a locality-based algorithm to solve this problem. One way of processing image locally is to cut the whole image plane into smaller blocks. For example, if an iris image is of size 30×360, and is cut into blocks of size 10×10, there will be $3 \times 36 = 108$ small blocks totally. Those blocks are called "patches." Processing image in patch-based algorithms has been widely used in variety of image processing field, such as texture analysis, edge and boundary detection, image segmentation, object recognition, biometric recognition, and generic image Super-Resolution.

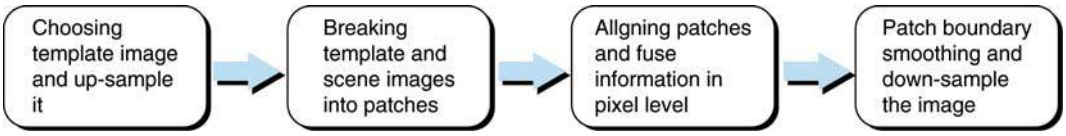
The patch-based Super-Resolution algorithm is quite intuitive. The first step is to cut every scene image into smaller patches. The second step is that for every location, align the local patch of every scene image with the template image. This step is important because iris images usually suffer from image deformation problem. This is especially true for segmented iris images since the pupil of an eye dilates when there is strong ambient lighting, and contracts when ambient lighting is weak. Therefore, it cannot be naively assumed that every patch from different images corresponds to exactly the same position on iris surface. Patches from different images have to be aligned, so that every pixel on the scene patch can be mapped correctly to the template image in the corresponding location.

After patches are locally aligned with each other, the third step is to combine information about numerical value for each pixel from scene patches and fill the holes on template patches with new value. If this process is illustrated with the metaphor of painting, this step is to draw details at the blank region on the template image. There are dozens of different algorithm to fuse information. One of the easiest ways is the method of linear combination. Suppose there are n numerical values from n different scene patches to fill one blank hole, the process of linear combination can be expressed as the following equation:

$$Y = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \sum_{k=1}^{n} \alpha_k x_k$$

where x_i is the numerical value of the ith scene patch, α_i is the coefficients of linear combination, and Y is the numerical value of the pixel after linear combination.

After all the blank holes are filled with the new values, the last step is to perform image smoothing across the boundaries of the patches. There are many existing smoothing algorithms. Linear, quadratic, or cubic interpolation can be used to achieve this goal. After the interpolation, the high-resolution image can be used or it can be down-sampled to make the final image the same size as the scene image but with much more detailed texture within. Figure 2 illustrates the flow chart of iris Super-Resolution algorithm.



Iris Super-Resolution. Figure 2 The flow chart of the iris Super-Resolution algorithm.

Summary

Image Super-Resolution is an important topic in image processing and has been extensively used in many applications, including quality enhancement for images and video, video surveillance, and biometric recognition. Super-Resolution for iris image is a very new topic. The technique can be applied to both iris synthesis and iris recognition. Not many detailed research results have been reported. In this section, one iris Super-Resolution algorithm is simply introduced by pixel-level information fusion from multiple low-resolution images. By performing Super-Resolution on iris image, the recognition rate of iris recognition system can be improved, and therefore, making the biometric system even more accurate and suitable for many real-life situations where high security issue is the top priority.

Related Entries

- ► Iris Image Quality
- ► Iris on the Move
- ► Iris Recognition, Overview
- ► Iris Sample Synthesis

References

1. Elad, M., Feuer, A.: Restoration of a single super-resolution image form several blurred, noisy, and under-sampled measured images. IEEE Trans. Image Process. **6**(12), 1626–1658 (1997)
2. Irani, M., Peleg, S.: Improving resolution by image registration. CVGIP: Graphics Models and Image Proc. **53**, 231–239 (1991)
3. Schulz, R.R., Stevenson, R.L.: Extraction of high resolution frames from video sequences. IEEE Trans. Image Process. **5**, 996–1011 (1996)
4. Hardie, R.C., Barnard, K.J., Armstrong, E.E.: Joint MAP registration and high resolution image estimation using a sequence of under-sampled images. IEEE Trans. Image Process. **6**, 1621–1633 (1997)
5. Farsiu, S., Robinson, D., Elad, M., Milanfar, P.: Advances and challenges in Super-Resolution. Int. J. Imaging Syst. Technol. **14**(2), 47–57 (2004)
6. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 1167–1183 (2000)
7. Capel, D.P., Zisserman, A.: Super-resolution from multiple views using learnt image models. Proceedings of IEEE International Conference, Computer Vision and Pattern Recognition **2**, 627–634 (2001)
8. Liu, C., Shum, H., Zhang, C.: A two step approach to hallucinating faces: global parametric model and local nonparametric model. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 192–198 (2001)
9. Freeman, W., Pasztor, E.: Learning low level vision. In: Proceedings of the Seventh International Conference on Computer Vision, 1182–1189 (1999)
10. Baker, S., Kanade, T.: Hallucinating faces. Proceedings of the International Conference on Automatic Face and Gesture Recognition, 83–88 (2000)

Iris Template Extraction Via Bit Inconsistency and GRIT

Gerry Vernon Dozier[1], Marios Savvides[2]
Kelvin Bryant[1], Taihei Munemoto[2]
Karl Ricanek, Jr.[3], Damon Woodard[4]
[1]North Carolina A&T State University
[2]Carnegie Mellon University
[3]University of North Carolina at Wilmington
[4]Clemson University

Synonyms

Bit fragility; Bit inconsistency; Fragile bits

Definition

The characteristic of iris code bits values being inconsistent (also referred to as *fragile*) across different images of the same iris was explored by Hollingsworth et al. [1]. The notion of fragile bits was first suggested

by Bolle et al. [2] when it was observed that the empirical false reject rate (FRR) was significantly better than predicted by their theoretical model. This fact implied that the bits of an iris code are not equally susceptible to "flip", given different environmental conditions that affect the quality of the captured iris images. Hollingsworth et al. demonstrated that by eliminating (masking) inconsistent bits, one could dramatically improve the FRR of an iris template.

Although the work of Hollingsworth et al. improves the FRR by identifying and removing fragile bits, our preliminary results show that it may be possible, based on bit instability, to further reduce the number of iris code bits needed for recognition. Our results show that this can also be done without an increase in the false accept rate (FAR). By using GRIT (Genetically Refined Iris Templates), iris code templates can be refined and transformed into templates that use a significantly smaller number of iris code bits for iris recognition. GRIT is a system that uses the concepts of bit inconsistency [1] and simulated evolution [3–6] in an effort to evolve iris code templates that use fewer iris code bits.

Bit Inconsistency

In [1], Hollingsworth et al. selected a dataset of 1251 images that were mostly unoccluded by eyelids or lashes. If an individual iris code bit was mostly one particular bit value and "flipped" to the other bit value some threshold percentage of the time, the bit was considered fragile (inconsistent). For their analysis, a bit was considered fragile if it flipped more than 40% of the time. Their results indicated that on average, 15% of the bits had probability greater than 40% of flipping and 85% of the bits had probability less than 40% of flipping. The implication of these rates indicates that the FRR of systems using Daugman-style [7] iris recognition systems could be dramatically reduced by focusing only on consistent bits. With this modified strategy, iris images would be analyzed to create an iris template that would mask out inconsistent bits. This mask would be combined with the typical masks used to eliminate eyelids and eyelashes.

Genetically Refined Iris Templates (GRIT)

GRIT is a system for evolving iris templates that have a decreased FRR; use a smaller number of iris code bits; and do not have an increased FAR. GRIT is composed of two components: a preprocessor and a Genetic Algorithm (GA) [3–6]. The GRIT preprocessor uses the concept of bit instability to eliminate fragile bits as well as develop a probability distribution function to be used by the GA to further reduce the number of iris code bits needed for recognition. GAs belong to a class of search techniques based on simulated evolution [5] and have been successfully used to solve a wide range of complex real-world search, optimization, and machine-learning problems.

The GRIT Preprocessor

In order to describe the GRIT preprocessor, let $I = \{i_0, i_1, \ldots, i_{n-1}\}$ and $M = \{m_0, m_1, \ldots, m_{n-1}\}$ be sets of iris codes along with their associated bit masks. The preprocessor, as shown in Fig. 1, works as follows. Given a set of iris codes and masks it sets $\imath$ to the first iris code, i_0, in the set of iris codes and μ to the associated mask of i_0, m_0. Next a hyper iris code and a hyper mask are created. The hyper iris code, v, simply records the number of iris codes in I that have the same value for each corresponding bit in $\imath$. The hyper mask, w, records the number of iris codes in I (along with their masks in M) that used a particular mask bit (associated with the best offset resulting in the best hamming ratio [7]) when comparing those iris codes with $\imath$, given μ.

Given v and w, the relative worth of flipping a bit in $\imath$ can be computed. The hyper gain, g, represents the number of iris codes in I that $\imath$ will become closer to

```
Procedure GRIT Preprocessor (I, M) {
  Let ι = i0 ∈ I and μ = m0 ∈ M;
  MakeHyperIrisCode (ι, I) → v;
  makeHyperMask (μ, M) → w;

  Let g = w – 2v;

  For j = 1 to # of Rows in Iris Codes
    For k = 1 to # of Columns in Iris Codes
      If (μjk == 1)
        If (gjk > 0) ι′ jk = ~ιjk;
        Else if (gjk == 0) μ′ jk = 0;

  Return (ι′, μ′, g)
}
```

Iris Template Extraction Via Bit Inconsistency and GRIT.

Figure 1 The GRIT Preprocessor.

(with respect to hamming ratio) by flipping a particular bit. Those bits in ι that have an associated positive gain can be flipped, because they will make the resulting iris template, ι', closer to a greater number of iris codes in I. Those bits in ι that have an associated gain of zero represent bits that have a 50% inconsistency rate, and therefore, can be "turned off'" (or masked) by setting the associated bit in μ to zero. The GRIT preprocessor returns a modified iris code, ι', and mask, μ', that has a lower FRR with respect to I than the iris template consisting of ι and μ. The hyper gain, g, is also returned and used as a probability distribution function (heuristic) to guide the mutation operator of the GRIT genetic search method. Figure 1 provides a pseudo-code example of the GRIT preprocessor.

The GRIT GA

The GRIT GA takes as input ι', μ', g, I, M, and δ, where δ represents the number of bits to mutate in creating an offspring template. The GRIT GA begins by creating P-1 mutants of μ', where P represents the population size of the GA. These mutants along with ι' make up the initial population of candidate bit masks. In creating the initial population, each of the mutants is created by mutating 100δ bits of μ' as follows. Two bit positions are selected at random and their associated gains (using g) are compared. The bit that has the highest gain is then mutated (flipped). This is how the Mutate method of the GRIT GA operates.

Each candidate bit mask in the initial population is evaluated using the following evaluation function:

$$F(\iota', c_{\text{mask}}, I, M, \tau) = \sum_j \beta(\iota', c_{\text{mask}}, \iota_j, \mu_j, \tau) + R(\iota', c_{\text{mask}}, I, M, \tau),$$

where c_{mask} represents a candidate mask, $\beta(\iota', c_{\text{mask}}, \iota_j, \mu_j, \tau)$ represents the number of iris code bits used in the comparison between the candidate iris template and the jth iris code in I, and $R(\iota', c_{\text{mask}}, I, M, \tau)$ represents the penalty if any of the iris codes in I is rejected by the template, $(\iota', c_{\text{mask}})$, by have a hamming ratio greater than a user-specified threshold, τ. The penalty is simply the summation of the amount by which the rejected iris codes exceed the threshold plus a constant value, α.

After each candidate mask has been evaluated and assigned a fitness by the evaluation function F, the counter, t, is set to |P| and the GA begins its evolutionary process. When the user-specified number of iris template evaluations has not been reached, the GA creates an offspring iris template mask by (1) selecting two individuals from the current population as parents; (2) crossing over [8] the parents to form an "embryo" by taking 50% of the bits from one parent and 50% of the bits from the other parent; and (3) mutating the "embryo". The offspring is then evaluated and replaces the worst-fit individual in the population (regardless of whether the worst-fit individual has a better fitness than the offspring). This process is repeated until the user-specified number of evaluations (Max_Evaluations) has been met. After the evolutionary process, the iris template with the lowest fitness in the population is then returned as the best solution evolved by the GA. Figure 2 provides a pseudo-code version of the GRIT GA.

The parent selection method works as follows: (1) randomly select two individuals from the population (excluding the worst-fit individual) and return the individual with the lower fitness as the first parent (mom); (2) repeat this process to get a second parent (dad). This parent selection method is commonly referred to as tournament selection [5].

Preliminary Results

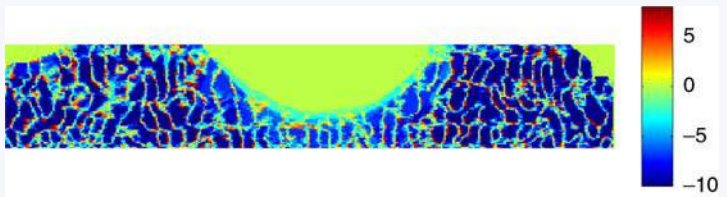
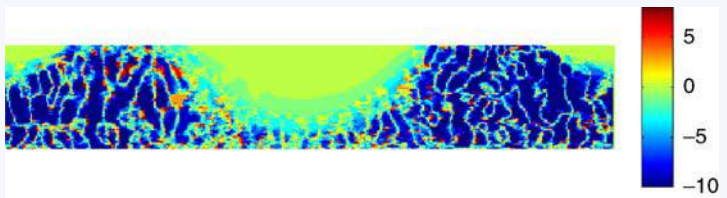
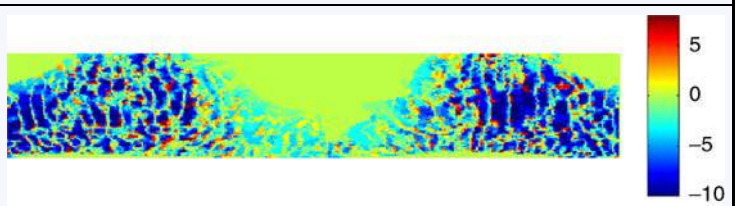
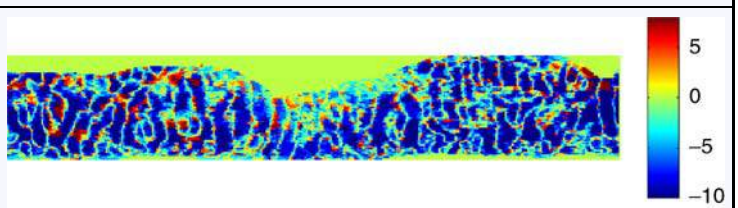
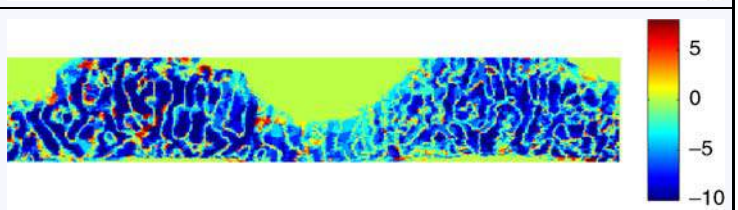
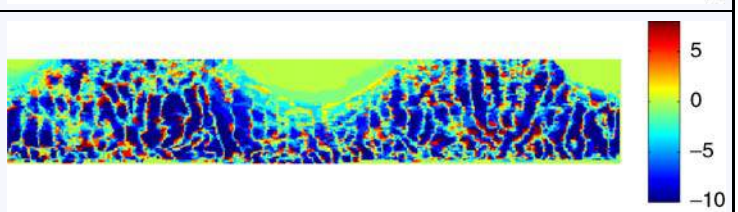
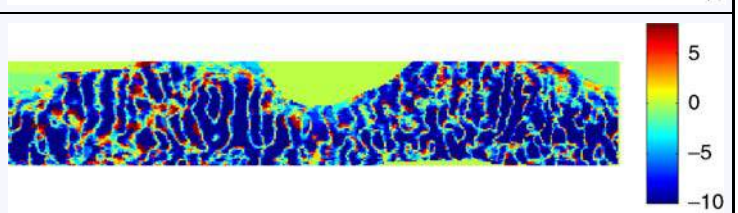
For this work, the iris images were segmented using the algorithm proposed by Thornton et al. [9], which finds

```
Procedure GRIT_GA (ι′, μ′, g, I, M, δ, τ) {
  InitalizePopulation (P_masks, μ′, g, 100δ);
  Evaluate (ι′, P_masks, I, M, τ);
  t = |P_masks|;

  While (t < Max_Evaluations) {
    wst  = GetWorstIndividual (P_masks);
    mom  = SelectParent (P_masks, wst);
    dad  = SelectParent (P_masks, wst);
    tmp  = Crossover (mom, dad);
    δ′   = δ [1.0- (t/Max_Evaluations)];
    kid  = Mutate (tmp, g, δ′);
    Evaluate (ι′, kid, I, M, τ);
    Replace (wst, kid);
    t = t + 1;
  }

  Return GetBestIndividual (P_masks);
}
```

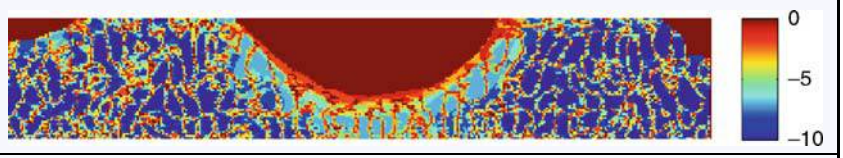
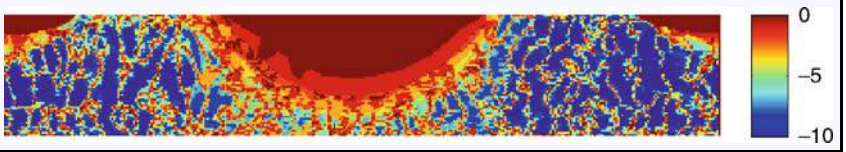
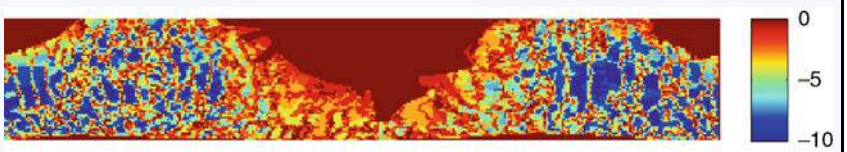
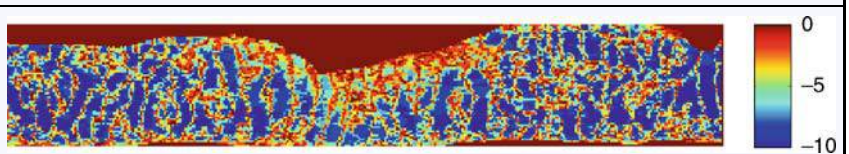
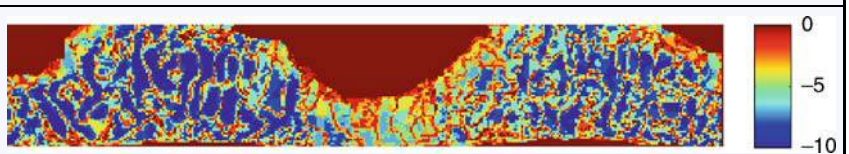
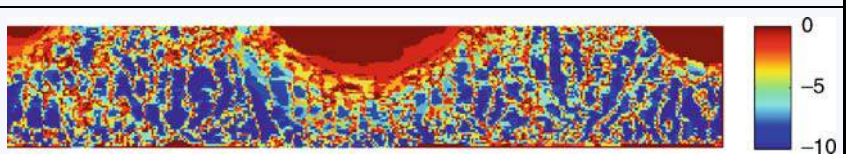
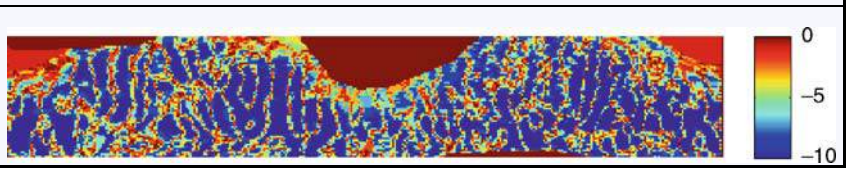
Iris Template Extraction Via Bit Inconsistency and GRIT.

Figure 2 GRIT (Genetically Refined Iris Templates).

Iris Template Extraction Via Bit Inconsistency and GRIT. Table 1 The Hyper Gains of Subjects Before Revising the Templates

Subject	Hyper Gain
4	
5	
38	
51	
67	
75	
93	

Iris Template Extraction Via Bit Inconsistency and GRIT. Table 2 The Hyper Gains of Subjects After Revising the Templates

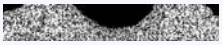
Subject	Hyper Gain
4	
5	
38	
51	
67	
75	
93	

nonconcentric circles for both pupil and iris boundaries with high accuracy. Eyelids and eyelashes were manually segmented. Given the iris images with occlusion masks, iris codes are generated by convolving them with a one-dimensional log Gabor filter row by row. As in the Daugman's algorithm [7, 10], the phase information at each pixel is quantized into two bits, and then Hamming distances are calculated by comparing these bits.

GRIT was used to develop iris templates for a total of seven subjects taken from the ICE 2006 dataset [11]: 4, 5, 38, 51, 67, 75, and 93. For these subjects, the first 10 iris codes in their respective sets were used to develop an iris template. The other 20 iris codes were used as a FRR test set. Each resulting iris template was checked with all of the other instances of the ICE 2006 dataset to determine its FAR.

The GRIT GA used a population size of 20 candidate bit masks and evolved an additional 4820 while keeping the best 19 candidate bit masks ever found at all times. The value of δ, the number of bits to mutate in creating an offspring, was set to 100, and the penalty

Iris Template Extraction Via Bit Inconsistency and GRIT. **Table 3** Preliminary GRIT Results

Subject	Mask Type	FRR	Bits Used	Visualized Masks
4	Original	0.0 (0.07)	31196	
	GRIT-PP	0.0 (0.07)	30545	
	GRIT-GA	0.0 (0.03)	20382	
5	Original	0.0 (0.00)	32250	
	GRIT-PP	0.0 (0.00)	31635	
	GRIT-GA	0.0 (0.00)	21200	
38	Original	0.6 (0.67)	30371	
	GRIT-PP	0.0 (0.00)	29667	
	GRIT-GA	0.0 (0.00)	19173	
51	Original	0.2 (0.13)	31456	
	GRIT-PP	0.0 (0.00)	30743	
	GRIT-GA	0.0 (0.00)	20169	
67	Original	0.1 (0.03)	31206	
	GRIT-PP	0.0 (0.00)	30455	
	GRIT-GA	0.0 (0.00)	19766	
75	Original	0.4 (0.37)	32724	
	GRIT-PP	0.0 (0.03)	31831	
	GRIT-GA	0.0 (0.03)	21759	
93	Original	0.0 (0.07)	33070	
	GRIT-PP	0.0 (0.00)	32365	
	GRIT-GA	0.0 (0.00)	21728	

constant, α, used in the penalty function, R, was set to 43,920, because the dimensions of the iris codes and masks used were $61 \times 360 \times 2$ bits.

Table 1 presents the hyper gains developed by the GRIT preprocessor for each of the seven subjects. In Table 1, for each visualization, the values associated with higher magnitudes (farthest away from zero) are more consistent (less fragile) than those bits with lower magnitudes. The red areas of a given subject represent the bits of the initial iris code template that were flipped in an effort to reduce the FRR of the 10 training instances. The green areas in the hyper gains represent those bits of the iris code mask that have been removed ("turned off"). The larger solid green regions that appear at the top of each of the hyper gains in Table 1 represent the initial masked bits. Table 2 presents the hyper gains after the iris template, $(\iota \ \mu)$, has been revised to form $(\iota' \ \mu')$. Notice that the visualizations show that there are a number of bits with hyper gain values close to zero that may be removed.

In Table 3, a comparison of the original masks, the masks developed by the GRIT preprocessor, and the

masks evolved by the GRIT GA are compared for each of the seven subjects. For each of the masks, the FRR on the training set of 10 instances is presented. For each of the masks, the number in parenthesis represents the FRR of the templates, compared with the 30 instances of I. In the next column, the average number of bits used in the comparisons with instances in the ICE dataset is presented. The final column in Table 3 shows a visualization of the original mask, the modified mask developed by the GRIT preprocessor, and the mask evolved by the GRIT GA. The FARs for all of the templates were zero.

In Table 3, notice that except for Subjects 4 and 75, the GRIT preprocessor was able to develop a modified iris code template and mask that reduced the FRR to zero. For Subject 4, the GRIT GA was able to reduce the FRR on the test set. This suggests that long runs of the GA may reduce the test set FRR further. Also, notice in Table 3 that the GRIT preprocessor is able to reduce the number of iris code bits needed to be used for recognition; however, the GRIT GA is able to reduce this number even further. The resulting iris codes bits needed are reduced by approximately 30% for each of the seven subjects.

Summary

In this article, GRIT, a novel approach toward developing iris templates, has been described. GRIT uses the concepts of bit inconsistency and genetic search to evolve iris templates that use a reduced number of iris code bits for iris recognition. The preliminary results show that the combination of bit inconsistency and genetic search provides a powerful hybrid for developing iris templates. Our results show that the reduction in the iris code bits needed does not result in an increase in FRR and FAR.

References

1. Hollingsworth, K., Bowyer, K., Flynn, P.: All iris code bits are not created equal. In: 2007 IEEE Conference on Biometrics: Theory, Applications, and Systems, September (2007)
2. Bolle, R.M., Pankanti, S., Connell, J.H., Ratha, N.: Iris individuality: A partial iris model. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 2, pp. 927–930 (2004)
3. Davis, L.: Handbook of Genetic Algorithms. New York, Van Nostrand Reinhold (1991)
4. Dozier, G., Homaifar, A., Tunstel, E., Battle, D.: An Introduction to Evolutionary Computation (Chapter 17). In: Zilouchian, A., Jamshidi, M. (eds.) Intelligent Control Systems Using Soft Computing Methodologies, pp. 365–380. Boca Raton, FL, CRC press (2001)
5. Fogel, D.B.: Evolutionary computation: Toward a new philosophy of machine intelligence, 2nd edn. Las Alomitas, IEEE Press (2000)
6. Goldberg, D.E.: Genetic Algorithms in Search, Optimization & Machine Learning. Addison-Wesley Publishing Company, Inc., Reading, Massachusetts (1989)
7. Daugman, J.: How iris recognition works. IEEE Trans. Circ. Syst. Video Technol. **14**(1), 21–30 (2004)
8. Syswerda, G.: Uniform Crossover in Genetic Algorithms. In: David S. (eds.) Proceedings of the Third International Conference on Genetic Algorithms (ICGA-89), pp. 2–9. San Francisco, CA, Morgan Kaufmann (1989)
9. Thornton, S.M., Kumar, V.: Robust iris recognition using advanced correlation techniques. Proceedings of the International Conference On Image Analysis and Recognition, pp. 1098–1105 (2005)
10. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. **15**(11), 1148–1161 (1993)
11. Iris Challenge Evaluation. National Institute of Standards and Technology, http://iris.nist.gov/ICE/, (2006)

Iris Template Protection

Patrizio Campisi, Emanuele Maiorana, Alessandro Neri
University of Roma TRE, Rome, Italy

Synonym

Iris template security

Definition

► Template protection is a crucial requirement when designing a biometric based authentication system. It refers to techniques used to make the stored template unaccessible to unauthorized users. From a template, information about the user can be revealed. Moreover, identity theft can occur. Therefore, it is of dramatic importance, if a template is compromised, to cancel, to revoke, or to renew it. Template protection can be performed using ► template distortion techniques,

► biometric cryptosystems, and ► data hiding techniques. Template protection techniques specifically designed and applied to ► iris images are hereafter summarized.

Introduction

Template protection is a key issue that has to be addressed when a biometric based authentication system is designed. It is highly desirable to keep secret a template both for security and for privacy reasons, and in case a template is compromised it is necessary to revoke, to cancel, or to renew it. Also, it is highly recommended to obtain from the same biometric different templates in order to avoid unauthorized tracking across different databases. In the recent past several techniques have been proposed to secure biometric templates and to provide the desirable cancelability and renewability properties. In the following limitations of classical cryptography, when applied within the biometric framework, are highlighted. Moreover, recently introduced techniques like template distortions, biometric cryptosystems, and data hiding techniques are briefly discussed first in general and later with specific application to iris template protection.

Cryptography [1] allows secure transmission of data over a reliable but insecure channel. The privacy of the message and its integrity are ensured, and the authenticity of the sender is guaranteed. However, cryptographic systems rely on the use of keys which must be stored and released on a password based authentication protocol. Therefore, the security of a cryptographic system relies on how robust is the password storage system to brute force attacks. However, template encryption cannot solve the biometric template protection problem. In fact, at the authentication stage, when a genuine biometrics is presented to the system, the match must be performed in the template domain, after decryption. However, this implies that there is no more security on the biometric templates. The match in the encrypted domain could solve this problem. However, because of the intrinsic noisy nature of biometric data, the match in the encrypted domain would inevitably bring to a failure because small differences between data would bring to significant differences between their encrypted versions. Some activities are flourishing to define signal processing operations in the encrypted domain, which could allow, for example, to perform operations on encrypted biometric templates on not trusted machines. However, this activity is still in its infancy and does not provide tools within the biometric framework yet.

Among the possible approaches recently proposed to address the issue of template protection, techniques based on intentional template distortions on the original biometrics have been introduced in [2]. Specifically, the distortion can take place either in the biometric domain, that is, before feature extraction or in the feature domain. Moroever, the distortion can be performed using either an invertible or a non invertible transform on the base of a user key which must be known at the authentication stage. Only the distorted data are stored in the database. This implies that, even if the database is compromised, the biometric data cannot be retrieved unless, when dealing with invertible transforms, user dependent keys are revealed. Moreover, different templates can be generated from the same original data, simply by changing the parameters of the employed transforms. The described technique allows obtaining both cancelability and renewability.

In the recent past, some efforts have been devoted to design *biometric cryptosystems* (see [3] for a review) where a classical password based authentication approach is replaced by biometric based authentication. Biometric cryptosystems can be used for either securing the keys obtained when using traditional cryptographic schemes or for providing the whole authentication system. A possible classification of the operating modes of a biometric cryptosystem is given in [3] where *key release*, *key binding*, and *key generation* modes are identified. Specifically, in the *key release* mode the cryptographic key is stored together with the biometric template and the other necessary information about the user. After a successful biometric matching, the key is released. However, this approach has several drawbacks, since it requires access to the stored template and then the one bit output of the biometric matcher can be overridden by using Trojan horse attacks. In the *key binding* mode the key is bound to the biometric template in such a way that both of them are inaccessible to an attacker and the key is released when a valid biometric is presented. It is worth pointing out that no match between the templates needs to be performed. Among the key binding approaches it is worth citing the fuzzy commitment

and the fuzzy vault scheme. In the *key generation* mode the key is obtained from the biometric data and no other user intervention besides the donation of the required biometrics is needed. Both the *key binding* and the *key generation* modes are more secure than the *key release* mode. However, they are more difficult to implement because of the variability of the biometric data.

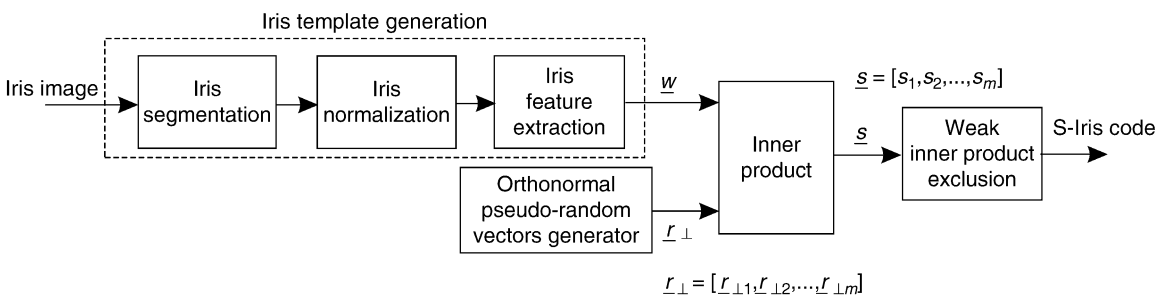
Data hiding techniques [4] complement encryption. In fact, encryption can be applied to ensure privacy, to protect the integrity, and to authenticate a biometric template. However, among the possible drawbacks, encryption does not provide any protection once the content is decrypted. On the other hand, data hiding techniques can be used to insert additional information, namely the watermark, into a digital object. Within the biometric framework, data hiding can be applied for copy protection, fingerprinting, data authentication, and timestamping in such a way that after the expiration date the template is useless. It is worth pointing out that some security requirements are also needed when dealing with data hiding techniques. In fact, according to the application, we should be able to face *unauthorized embedding*, *unauthorized extraction*, and *unauthorized removal* of the watermark. Recently some efforts are being devoted to the integration between watermarking and cryptography. However, much more research activity is still needed before deployment. In the following, after a quick overview on iris template generation, the most significant approaches for iris template protection are discribed.

Iris Template Generation

An iris image is preprocessed to select the actual iris region to use for feature extraction, thus removing unwanted elements such as eyelid, eyelashes, pupil, reflections, and all the other noise components. Then an iris normalization process takes place, since the extracted iris regions, both from different people and from the same people, can differ because illumination changes, variation of the eye-camera distance, elastic deformations in the iris texture, and similar. These effects can generate matching problems. In some approaches a scale-invariant transform like the Fourier-Mellin is used. In some others, a mapping of the iris image from raw cartesian coordinates to non concentric polar coordinate system is used. After the normalization stage, the features extraction procedure takes place. This task can be accomplished using different approaches such as multi-scale Gabor wavelet filtering and its variants, singular value decomposition, principal component analysis, and so on.

Cancelable Iris Template

A cancelable iris biometric approach, namely S-Iris Encoding, is proposed in [5]. The method is roughly sketched in Fig. 1 and briefly summarized in the following. Iris preprocessing is performed first. Specifically iris segmentation by means of the Canny edge detector, to find the edge map, followed by the Circular Hough Transform, to detect the iris and pupils boundaries are carried out. Linear Hough transform is used to discard eyelids and eyelashes. The normalization is performed using the Daugman's rubber sheet model [6]. The iris feature extraction is performed by convolving the normalized 2D pattern rows, each corresponding to a circular ring of the iris region, by using 1D Log-Gabor filter. The magnitude of the so obtained complex features are then collected in a vector $\underline{w}$ that is further processed to obtain the S-Iris code



Iris Template Protection. Figure 1 S-Iris Encoding scheme [5].

as described in the next steps. A set of m orthonormal pseudorandom vectors $\{\underline{r}_{\perp,i}\}$, with $i = 1, 2, \cdots, m$, are then generated using a token. The inner products $\alpha_i = <\underline{w}, \underline{r}_{\perp,i}>$, with $i = 1, 2, \cdots, m$, are then evaluated. The m bits of the S-Iris code $\underline{s} = \{s_i \mid i = 1, \cdots, m\}$ are computed as

$$s_i = \begin{cases} 0 & \text{if} \quad \alpha_i < \mu_i - \sigma_i, \ \alpha_i > \mu_i + \sigma_i \\ 1 & \text{if} \quad \mu_i - \sigma_i \leq \alpha_i \leq \mu_i + \sigma_i, \end{cases}$$

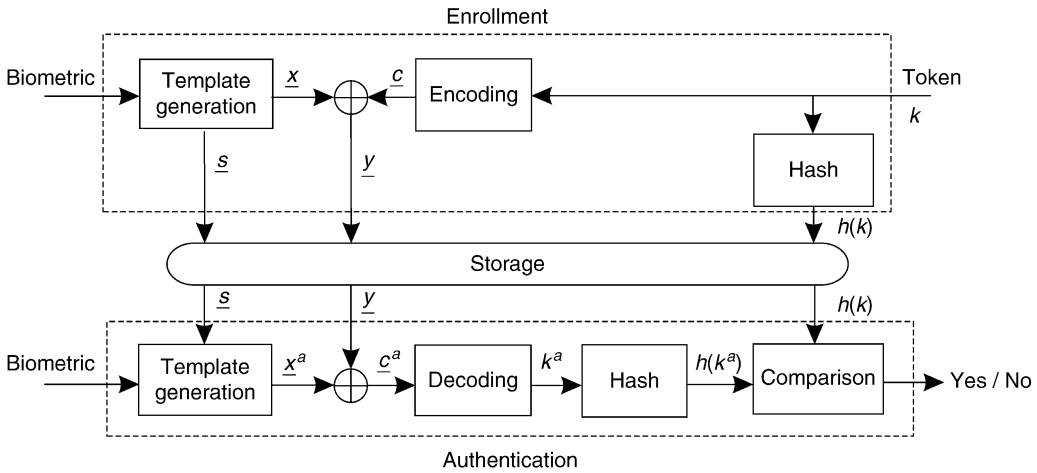
where μ_i and σ_i are the average and standard deviation of α_i respectively. This approach allows to discard those inner products which are numerically small and which therefore must be excluded in order to improve the verification rate. The authors of [5] point out that the system authentication performance have a significant improvement over the solely biometric system.

Iris Template Protection using Cryptosystems

Among the methods which can be classified as key binding based approaches [3] we can cite the fuzzy commitment scheme [7], based on the use of error correction codes and the fuzzy vault scheme [8], based on polynomial based secret sharing.

Specifically, the fuzzy commitment scheme is depicted in Fig. 2 in its general form. In the enrollment stage, the biometric template $\underline{x}$ is used to derive some side information $\underline{s}$ which is stored to be used in the authentication stage. Then a randomly chosen codeword $\underline{c}$ is generated on the base of a token k. The binding between the biometric measurement $\underline{x}$ and the codeword $\underline{c}$ is obtained as $\underline{y} = \underline{x} \oplus \underline{c}$. Both $\underline{y}$ and a hashed version of the token k are eventually stored. In the authentication stage, the side information $\underline{s}$ is retrieved and, together with the actual biometric measurement, it is used to obtain the biometric template $\underline{x}^a$. This latter usually differs from the template obtained in the enrollment stage because of the intrinsic variability of biometrics. Then the codeword $\underline{c}^a$ is obtained as $\underline{c}^a = \underline{x}^a \oplus \underline{y}$. Finally k^a is obtained by decoding $\underline{x}^a$. Its hashed version $h(k^a)$ is obtained and compared with the stored $h(k)$. If the obtained values are identical, the authentication is successful. It is worth pointing out that this scheme provides both template protection, since from the stored information $(\underline{s}, \underline{y}, h(k))$ it is not possible to retrieve the template, and template renewability, since by changing the token k the template representation changes.

In [9] the fuzzy commitment scheme here described is applied to iris protection. Iris preprocessing consists in the edge map extraction followed by Circular Hough Transform to detect the iris and pupils boundaries followed by Linear Hough transform to discard eyelids and eyelashes. The normalization is performed using the Daugman's rubber sheet model. The iris feature extraction is performed by convolving the rows of the normalized 2D pattern by using 1D Log-Gabor filter. The phase information from both the real and the imaginary part is eventually quantized. A reliable bits selection is then performed according to the assumption that the more reliable bits are those coming from the pixels closer to the pupil center,



Iris Template Protection. **Figure 2** Fuzzy Commitment scheme.

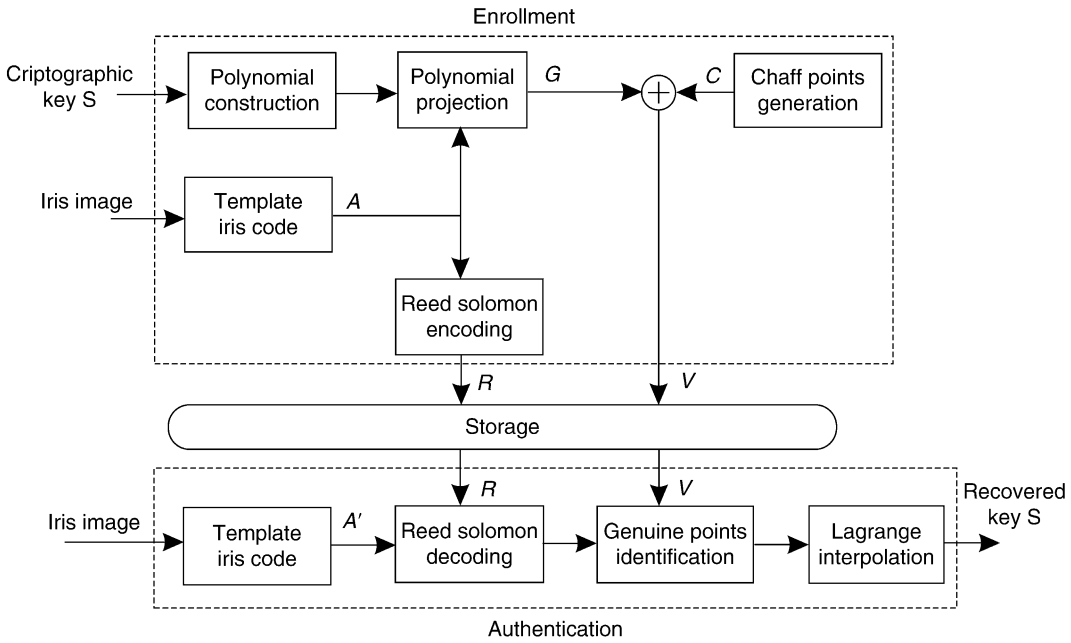
where eyelid and eyelashes are not likely to be found. With respect to the general scheme in Figure 2, in [9], the feature vector $\underline{x}$ is split into two feature vectors $\underline{x}_1$ and $\underline{x}_2$ of the same length and two BCH encoder are used. Specifically, two tokens k_1 and k_2 are employed to generate two codewords $\underline{c}_1$ and $\underline{c}_2$ obtained each from one of the two employed BCH encoders. Eventually the secret data $\underline{y}_1 = \underline{x}_1 \oplus \underline{c}_1$ and $\underline{y}_2 = \underline{x}_2 \oplus \underline{c}_2$ are obtained. Therefore the stored information will be given by $(\underline{s}, \underline{y}_1, \underline{y}_2, h(k_1), h(k_2))$. The authentication step is dual with respect to the enrolment stage. The authors of [9] point out that the division strategy is needed to balance the desired verification accuracy and the BCH code error correction capability.

In [10] the authors use the fuzzy commitment scheme for the protection of binary iris template, namely the iriscode [6], by employing a cascade of Reed Solomon codes and Hadamard codes to handle the intra-variability of the biometric templates. This choice has been driven by an exhaustive study of the error patterns which can be encountered employing iris codes. The authors propose a fuzzy commitment architecture, where a two-layer error correction method is performed. The outer layer uses a Hadamard code to correct random errors at the binary level which are generated by CCD camera pixel noise, iris distortion, or other image-capture effects which can not be corrected by the initial preprocessing. The inner layer uses a Reed Solomon code to correct burst errors in the iriscode, due to undetected artefact like eyelashes or specular reflections in the iris image. The proposed architecture is tested on a proprietary database with 700 iris samples from 70 different eyes, with 10 samples from each eye. It has been found out that an error free key can be reproduced from an actual iriscode with a 99.5% success rate. Iris orientation is of big concern when unlocking the key in the fuzzy commitment scheme. Multiple attempts have to be performed, shifting the observed iris code by octect-bits, being impossible to cyclically scroll the iris sample as in the unprotected approach.

In [11] the application of the fuzzy commitment scheme, for the protection of biometric data, is discussed. Specifically, a method for finding an upper bound on the underlying error correction capability, when using a fuzzy commitment scheme is provided. The analysis is conducted by introducing a model for the recognition process, composed of two binary symmetric channels, the matching and the non matching channel. Specifically, the first is used to model the errors coming from the matching between templates belonging to the same user. The latter is used to model the errors coming from the matching between templates belonging to different users. An erasure mechanism is introduced in the matching channel to manage the template dimension variability due for example to occlusions. Moreover, a practical implementation of the fuzzy commitment for iris template protection is proposed, employing as error correcting codes the product of two Reed Muller codes, together with a specific decoding process, derived from the min-sum decoding algorithm. The proposed protection scheme is tested on a public iris database. The authors show that correction performance close to the theoretical optimal decoding rate are obtained.

The fuzzy vault cryptographic scheme [8] consists in placing a secret S in a vault and in securing it by using a set of unordered data $A = \{a_1, a_2, \cdots, a_N\}$, which in our biometric context represents the biometric template. Specifically, a polynomial $p(x)$, whose coefficients are given by the secret S, is generated and the polynomial projections $p(a_i)$, for all the elements belonging to A, are evaluated. Then a large number of chaff points, which do not lie on the polynomial $p(x)$, are arbitarily chosen. Specifically, M unique points $\{c_1, c_2, \cdots, c_M\}$ are randomly set with the constraint that $c_j \neq a_i$, for $j = 1, 2, \cdots, M$ and $i = 1, 2, \cdots, N$. Then, another set of M random points $\{d_1, d_2, \cdots, d_M\}$, such that $d_j \neq p(c_j)$, $j = 1, 2, \cdots, M$, is chosen. The concatenation of the two sets $\{(a_1, p(a_1)), (a_2, p(a_2)), \cdots, (a_N, p(a_N))\}$ and $\{(c_1, d_1), (c_2, d_2), \cdots, (c_M, d_M)\}$ represents the vault V which secures both the secret and the template. When a user tries to unlock the vault, another set of unordered data A' can be used. If the set A' substantially overlaps with the set A then the user can identify many points of the vault lying on the polynomial. If the overlapping point number is sufficient, the polynomial can be identified by using Lagrange interpolation, thus identifying the secret. If the two sets are significantly different, the polynomial reconstruction is unfeasible. Many implementations of the general principle here sketched have been proposed in literature.

In [12], iris data are used for securing the vault. The method is depicted in Fig. 3. Specifically, the feature extraction is performed as follows. After having localized the iris region, it is transformed into a polar coordinate image and two regions not occluded by



Iris Template Protection. **Figure 3** Fuzzy Vault scheme as in [12].

eyelids and eyelashes are selected. From each selected region, eight iris blocks are derived and transformed using Independent Component Analysis thus obtaining 16 feature vectors. In order to take into account the intra-class variations, the blocks extracted from each image are clustered, employing a K-means algorithm, thus generating an iris code of sixteen 8-bit symbols, which represents the elements of the locking set A. The vault locking is performed as sketched in Figure 3 and uses the general principle of the fuzzy vault scheme. However, in the implementation proposed in [12], the locking set A is also encoded using Reed Solomon codes, thus obtaining a set R which is stored together with the set V obtained by concatenating the genuine points at v obtained concatenation $G = \{(a_1, p(a_1)), (a_2, p(a_2)), \cdots, (a_N, p(a_N))\}$, coming from the polynomial $p(x)$, and the set c derived from the chaff point set $C = \{(c_1, d_1), (c_2, d_2), \cdots, (c_M, d_M)\}$. The set R is employed during authentication to correct potential errors, due to intra-class variations, in the query iriscode. When the query iris image is analyzed during authentication, the iris blocks are extracted, compared with the cluster maps thus generating a new iris code, which is corrected using the stored Reed Solomon redundancy set and employed to unlock the vault. The secret key is thus revealed.

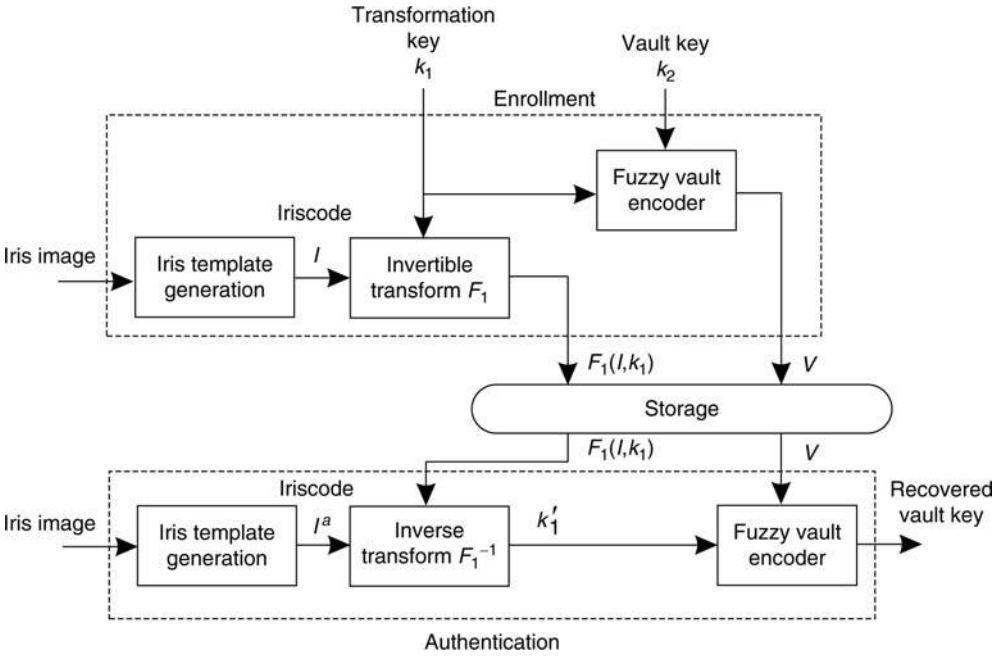
In [13] an iris cryptosystem relying on an invertible transform, followed by fuzzy vault locking has been proposed to secure the iriscode [6]. The scheme is given in Fig. 4. More in detail the proposed iris cryptosystem is a two step process. In the first step, an invertible transform F_1, chosen on the base of a randomly generated transformation key k_1, is applied to the iriscode I. In the second step, the fuzzy vault scheme, with key k_2, is applied to secure the transformation key k_1, thus giving the vault V. Both the transformed iriscode and the vault, which locks the transformation key k_1, are eventually stored. In the authentication stage, the inverse transformation F_1^{-1} is applied to the transformed iris code template using the query iriscode I^a, thus obtaining the transformation key $k_1^{'}$. Then the transformation key $k_1^{'}$ is used to decode the vault V. If the vault key k_2 is successfully recovered this implies that there is a match between the iriscode template I and the iriscode used in the authentication stage I^a. The author of [13] points out that both the invertible transform and the fuzzy vault introduce error correction, therefore the proposed cryptosystem is able to manage a higher intra class variation.

The protection of iris templates is also discussed in [14], where a trade-off between the authentication performances and the security of key binding schemes, is discussed from an information theoretic perspective. A practical cryptosystem for iris templates, based on Low Density Parity Codes (LDPC) and belief propagation, is also proposed. The pre-processing of the iris

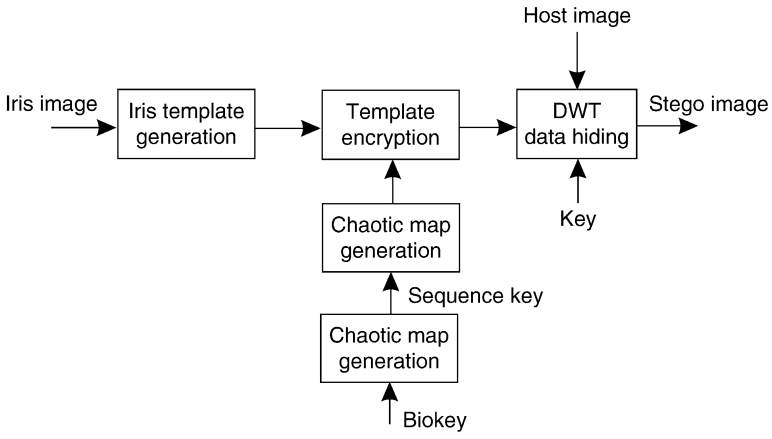
images is performed according to [6], thus obtaining a binary sequence $\underline{m}$. The bits corresponding to unreliable positions, identified during training, are discarded from m. The obtained binary vector $\underline{z}$ is then mapped into the secure biometric S by computing the syndrome of $\underline{z}$ with respect to a low density parity check code, whose parity check matrix H is randomly chosen. When a user claims his identity, the realiable feature vector $\underline{z}'$ is computed and a belief propagation algorithm is applied to retrieve the sequence whose syndrome is S. The trade-off between the False Rejection Rate and the security of the proposed implementation is discussed, while the False Acceptance Rate is not taken into account.

Iris Template Protection using Data Hiding

In [15] a steganagraphic technique for covert communication of biometric data using chaos theory is proposed with application to irisdata. The proposed data hiding scheme is sketched in Fig. 5. Specifically, the iris template, namely the iriscode, is extracted using the



Iris Template Protection. Figure 4 Iris cryptosystem [13].



Iris Template Protection. Figure 5 Chaos based data hiding scheme [15].

method in [6]. Then two chaotic maps are used for encrypting the iris template. The first map is used to generate a 1D sequence of real numbers used as a sequence key. A biometric generated key, the biokey, is used to set the initial condition and the parameters of the chaotic map. Then, the so obtained 1D sequence is used as the sequence key of a different chaotic map which is used to encrypt the template. The authors of [15] point out that this approach assures robustness against different kind of attacks. After encryption, the template is embedded into the cover image by using a discrete wavelet transform (DWT) decomposition. The template extraction and decryption is made on the authentication side by performing dual operations with respect to the ones done at the embedding side. The authors highlights that their method offers better performance than those given by using only one chaotic map.

Summary

Template protection is a key requirement when designing a biometric based authentication system. A brief overview of the main approaches based on the use of transforms, biometric cryptosystems, and data hiding techniques, either specifically tailored or simply applied to iris template protection have been here outlined.

Related Entries

- ▶ Biometric Encryption
- ▶ Biometric Security Overview
- ▶ Iris Databases
- ▶ Iris Digital Watermarking
- ▶ Template Security

References

1. Menezes, A., van Oorschot, P., Vanstone, S.: Handbook of applied cryptography, CRC Press (1996)
2. Bolle, R.M., Connell, J.H., Ratha, N.K.: Biometric perils and patches. Pattern Recognit. **35**, 2727–2738 (2002)
3. Jain, A.K., Nandakumar, K., Nagar, A.: Biometric template security. EURASIP J. Adv. Signal. Process, Special Issue on Advanced signal processing and pattern recognition methods for Biometrics, article ID 579416 (2008)
4. Cox, I., Miller, M., Bloom, J., Miller, M., Fridrich, J.: Digital watermarking and steganography 2nd edn. Morgan Kaufmann (2007)
5. Chin, C.S., Teoh, A.B.J., Ngo, D.C.L.: High security iris verification system based on random secret integration. Comput. Vis. Image Underst. **102**(2), 169–177 (2006)
6. Daugman, J.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 21–30 (2004)
7. Juels, A., Wattenberg, M.: A fuzzy commitment scheme, In: Sixth ACM Conference on Computer and Communication Security, pp. 28–36 (1999)
8. Juels, A., Sudan, M.: A fuzzy vault scheme, In: Proceedings of the IEEE on International Symposium on Information Theory, pp. 408 (2002)
9. Yang, S., Verbauwhede, I.: Secure iris verification, In: Proceedings of the IEEE ICASSP 2007, vol. 2, pp. 133–136 (2007)
10. Hao, F., Anderson, R., Daugman, J.: Combining crypto with biometrics effectively. IEEE Trans. Comput. **55**(9), 1081–1088 (2006)
11. Bringer, J., Chabanne, H., Cohen, G., Kindarji, B., Zemor, G.: Optimal iris fuzzy sketches. In: First IEEE International Conference on Biometrics: Theory, Applications, and Systems, BTAS 2007, pp. 1–6 (2007)
12. Lee, Y.J., Bae, K., Lee, S.J., Park, K.R.: Biometric key binding: Fuzzy vault based on iris images. ICB 2007, LNCS 4642, pp. 800–808 (2002)
13. Nandakumar, K.: Multibiometric systems: Fusion strategies and template security, Dissertation, Michigan State University, Department of Computer Science and Engineering (2008)
14. Martinian, E., Yekhanin, S., Yedidia, J.S.: Secure biometrics via syndromes. In: 43rd Annual Allerton Conference on Communications, Control, and Computing, Monticello, IL, Oct (2005)
15. Khan, M.K., Zhang, J., Tian, L.: Chaotic secure content-based hidden transmission of biometric templates, Chaos Solitons & Fractals, Elsevier, 32, pp. 1749–1759 (2007)

Iris Template Security

▶ Iris Template Protection

Iris2pi

This is a most widely used iris recognition algorithm as of 2008. This is a version of the Daugman algorithm. It differs from an earlier version, "bowtie," in the way it handles eyelid (and other) occlusions. The bowtie

algorithm analyzed a bowtie shaped section of the iris – two triangular wedges extending to the left and right – that avoided most eyelid occlusion at the expense of throwing away information for eyes that are wide open. Iris2pi attempts to analyze all the iris that can be seen – a full 2π radians if possible. Iris2pi note regions that are occluded or otherwise invalid for biometric identification; it records that information in the biometric template. When two templates are compared, the algorithm only compares regions that have valid data.

► Iris Device
► Iris Encoding and Recognition using Gabor Wavelets

IrisCode

IrisCode is a digitized, normalized, compact encoding of the unique texture visible in the iris of an eye, for purposes of automated biometric identification. The IrisCode is mapped between the inner and outer boundaries of the iris, so it is size-invariant, distance-invariant, and also invariant to changes in pupil dilation. This intrinsic normalization facilitates the searching and matching operations. In the standard format (called "iris2pi") used in public deployments of iris recognition, the IrisCode is based on a phase encoding by Gabor wavelets, and it also incorporates masking bits signifying the detection of eyelids, eyelashes, reflections, or other noise. Standard code lengths are 512 or 1,024 bytes. The IrisCode enables simple parallel logical operators XOR (Exclusive-OR) and AND to generate Hamming Distance scores for similarity between IrisCodes, at speeds of typically 1 million complete IrisCode comparisons per second.

► Iris Encoding and Recognition using Gabor Wavelets
► Iris Recognition at Airports and Border-Crossings
► Score Normalization Rules in Iris Recognition
► Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition

ISO

ISO is an acronym referring to International Standard Organization, an international entity responsible for defining standards and providing certifications of compliance.

► Biometric Sensor and Device, Overview

J

JPEG and JPEG2000 Image Compression

Images can be encoded much more efficiently than by pixel arrays if local regions are represented as combinations of elementary functions. Computing the coefficients on those elementary functions such that their linear combination becomes equivalent to, or closely approximates, the original image is the same operation as computing a transform. Each local image region is multiplied by each of several such elementary functions and integrated to obtain each such coefficient. The resulting coefficients usually have lower entropy than the original pixel distribution, enabling more compact coding; in addition, their values can be coarsely quantized without detrimental effect. An image is recovered from the coded coefficients by essentially an inverse transform. The most ubiquitous image compression protocol is JPEG, defined by ISO Standard 10918. It applies the Discrete Cosine Transform (DCT) to local square tiles of an image (typically 8×8 pixels), but the abrupt truncation of each cosine wave causes "block quantization" artifacts which become noticeable when only subsets of cosine waves are used in order to achieve compression ratios above about 30:1. JPEG2000 overcomes this problem by replacing the block DCT cosine waves with Daubechies wavelets which are smoothly attenuated instead of chopped; the resulting Discrete Wavelet Transform (DWT) is the core of JPEG2000 ISO Standard 15444. JPEG2000 also has other advanced features to allocate the coding budget inhomogeneously across an image if needed. Both protocols allow control over the compression factor (CF for JPEG2000; quality factor QF for JPEG). Despite its superior mathematical basis and performance, JPEG2000 is not as widely used as JPEG nor as freely available.

► Iris Recognition Performance Under Extreme Image Compression

© 2009 Springer Science+Business Media, LLC

K

Kernel

A kernel k is a function that for all $\mathbf{x}, \mathbf{z} \in \mathbb{X}$: satisfies $k(\mathbf{x}, \mathbf{z}) = \Phi(\mathbf{x}) \cdot \Phi(\mathbf{z})$, where Φ is a mapping from the input space $\mathbb{X}$ to the feature space $\mathcal{H}$, i.e., $\Phi : x \mapsto \Phi(x) \in \mathcal{H}$. A kernel function can also be characterized as follows: Let $\mathbb{X}$ be the input space. A function k: $\mathbb{X} \times \mathbb{X}$: $\mapsto \mathbb{R}$ (or $\mathbb{C}$) is kernel *if and only if* for any $M \in \mathbb{N}$ and any finite data set $\{\mathbf{x}_1, \cdots, \mathbf{x}_M\} \subset \mathbb{X}$, the associated Gram matrix is positive semi-definite.

▶ Non-linear Techniques for Dimension Reduction

Key Binding

▶ Biometric and User Data, Binding of

Keypoints

▶ Local Image Features

Keystroke Dynamics

▶ Keystroke Recognition

Keystroke Pattern Classification

▶ Keystroke Recognition

Keystroke Recognition

Nick Bartlow
West Virginia University, Morgantown, WV, USA

Synonyms

Behavioral biometrics; Keystroke dynamics; Keystroke pattern classification

Definition

Keystroke recognition is a ▶behavioral biometric which utilizes the unique manner in which a person types to verify the identity of an individual. Typing patterns are predominantly extracted from computer keyboards, but the information can potentially be gathered from any input device having traditional keys with tactile response (i.e., cellular phones, PDA's, etc). Although other measurements are conceivable, patterns used in keystroke dynamics are derived mainly from the two events that make up a keystroke: the Key-Down and Key-Up. The Key-Down event takes place at the initial depression of a key and the Key-Up occurs at the subsequent release of that key. Various unique features are then calculated based on the intra-key and inter-key timing variations between these events. After feature extraction, a wide range of algorithms can be employed to establish whether the unique pattern confirms or denies the claimed identity.

© 2009 Springer Science+Business Media, LLC

Introduction

The earliest form of keystroke recognition emerged in the early 1900s during the days of World War I. During the war, the French used listening posts in which operators were able to recognize the "fist" of enemy radio operators communicating in Morse code. These trained individuals would learn to recognize operators by differing lengths of pauses, dots and slashes, and varying transmission speeds. This intelligence subsequently allowed the French to establish the identity of entities such as enemy battalions. Far more sophisticated than electromechanical telegraphs used to transmit Morse code, keyboards of today offer many more opportunities to establish the unique manner in which one types. Intuitively, coarse level differentiation can be achieved by investigating typing speeds. For instance, a professional typist who averages 90 or more words per minute would be easily distinguished from a "hunt and peck" amateur who averages only 20–25 words per minute. That said, this feature only goes so far as many people type at similar speeds and the average speed that an individual types can vary significantly depending on many factors. The time it takes an individual to locate a key (sometimes referred to as "seek-time") also varies from key to key. For instance, left-handed individuals may have quicker seek-times for keys on the left side of the keyboard and vice versa [1]. Along those same lines, use of the shift keys to modify characters can also vary from individual based on handedness and typing skill. Trained professionals will always modify characters on the right side of the keyboard with the left shift key while amateurs may continually use the right shift key to do so [2]. Language undoubtedly plays a large role in the individuality of a typing signature. Given that a person speaks English, commonly used words like {the, and, you, are} are often "programmed" in one's mind and typed quickly as opposed to an individual of a different native language. Additionally, individuals typically exhibit a consistent pattern of errors including replacements, reversals, and extraneous hits. In an extreme case, the consistent lack of errors is a pattern in itself.

Keyboard Technology and Semantics

There are four different kinds of switch technology used in keyboards today; pure mechanical, foam element, rubber dome, and membrane [3]. Each switch type has various characteristics such as feel, durability, price, etc. No matter the key switch technology chosen, when a key is depressed, a degree of "bounce" is present. Bounce can be defined as the effect when the contact device rapidly engages and disengages over an extremely short period of time [3]. Keyboards, either external to desktop PCs or internal to laptops and other devices are computers in their own right as they contain a microprocessor, RAM, and sometimes ROM. Using their processors and controllers, they filter out the difference between bounce and two successive keystrokes. Each stroke therefore consists of two events, when the plates are engaged and when the engagement is released or disengaged. Scan codes resulting from these events are sent from the controller in the keyboard to the event handler in the BIOS of the device in question (usually a PC) [3]. Scan codes are recorded by the processor based on a matrix composed of all the keys on the keyboard. The keyboard matrix operates on a buffer that allows for the processing of simultaneous keystroke events. As mentioned before, when a key is pressed down, the plates become engaged. It is at this point that the keyboard processor sends a "make code" encoded as a hex value to the device. The make code can be thought of as including both the key engaged and various other state flags indicating if/how the key was modified by any of the various control keys such as shift, alt, etc. Once the key disengages, a corresponding "break code" is sent to the PC [3]. These ideas form the basis of keyboard technology at its lowest-level.

Using this background as a foundation, the upper level semantics of keyboard operation can be defined. The basis of all features included in keystroke recognition is founded on the keystroke event and the associated make code or break code correlation described previously. Instead of dealing with terms like "make code," "disengagement," etc., researchers usually yield to the more intuitive, higher level definitions below.

1. *Key-down.* The event that fires when a key is pressed down. This corresponds to the event of the keyboard processor sending the device (usually a PC) a "make code." It should be noted that this event will continually fire until the key being depressed is released. The speed at which the Key-down event fires while a key is depressed is referred as the "repeat rate." This is a user customizable property in virtually all operating systems.

2. *Key-up.* The event that fires when a currently depressed key is subsequently released.
3. *Keystroke.* The combination of an initial Key-down event and the corresponding Key-up event.
4. *Hold time.* The length of time between an initial Key-down event and the corresponding Key-up event. Hold time is sometimes referred to as "dwell time."
5. *Delay.* The length of time between two successive keystrokes. It should be noted that this time can be positive or negative (overlapping strokes). Some works refer to delay as "latency" or "flight."

Some highly specialized keyboards can record other information such as the pressure of key strikes, but the foundation of the technology is based on the events defined above.

Feature Representation and Classification

A wide variety of algorithmic approaches have been explored as suitable candidates for the task of keystroke recognition. The problem of keystroke recognition fits well within the general fields of pattern recognition and machine learning; the two main tasks involved in solving problems within these fields are to define the representation of the feature space and the algorithm used to predict the class of samples. As mentioned in previous sections, the features in keystroke recognition are primarily derived from the elements that make up a keystroke. Most algorithms utilize first order statistics such as minimum, maximum, mean, median, and standard deviation of hold times and latencies [2, 4–8] for feature representation. Here, hold times are for individual keys whereas latencies are measured between two keystrokes often defined as "digraphs." Using these statistics, one can either calculate fixed length feature vectors as outlined in [2] or variable length feature vectors as outline in [9]. Fixed length or static size feature vectors will always have a predetermined length despite the length of the input sequence. The size of variable length or of dynamic feature vectors will depend on the size of the input sequence. Although the vast majority of keystroke recognition systems rely on single key hold times and digraph latencies, some approaches define other feature sets including trigraph durations, ordering of keystrokes (when shift-key modification is required), etc. [9].

Beyond feature representation, a keystroke recognition system must employ an algorithm to predict the class of incoming samples. In general, the approaches can be broken down into two sections: distance metric based approaches and machine learning approaches. After calculating the feature vector for an incoming sample, the chosen algorithm must predict the class of the sample (genuine or imposter). Many approaches will do so by comparing the incoming sample to one or more reference samples in a template database through a distance metric. Popular distance metrics include: Euclidean, Mahalanobis, Manhattan, Chebyshev, and Hamming. When distance metrics are employed to compare two samples, the smaller the score the closer the two samples are to each other. Gaines and Lisowski [4], Garcia [10], Young and Hammon [11], and Joyce and Gupta [5] are all examples of algorithms that utilize one or more of these distance metrics as classification schemes. Table 1 provides an overview of selected work in keystroke recognition including the works listed above. The table includes the features/algorithm used, input requirements, the scope, and performance. Under the performance column the raw totals in terms of FAR and FRR are presented within parentheses when listed in the work.

As the field has matured, many other machine learning approaches have emerged as viable solutions for prediction mechanisms in keystroke recognition. Neural networks have widely been employed with works by Obaidat et al. [6, 7], Brown et al. [12], and Maisuria et al. [13]. Cho and Yu have applied Support Vector Machines (SVM's) to the problem extensively [14, 15]. Additionally, Bartlow and Cukic explored the decision tree approach of Random Forests [2] (see Table 1 for more information on listed works).

Applications and Challenges

In application, the uses of keystroke recognition can range anywhere from stand-alone biometric systems to augmenting general computer security systems. Depending on various system specific security characteristics such as database size and operational risks, keystroke recognition is suitable as a stand-alone biometric. Although not on the level of physiological biometrics such as iris, fingerprint, and face, many works in the literature indicate that the attainable performance rates are within the scope of what some operational profiles would require. Much like the physiological biometrics,

Keystroke Recognition. Table 1 Overview of Selected Works in Keystroke Recognition

Work	Feature(s)/Algorithm	Input	Scope	Performance
Gaines and Lisowski (1980) [4]	Latency between 87 lowercase digraphs using sample t-tests	300–400 word passage 2 times	7 secretaries	FAR 0% (0/55) FRR 4% (2/55)
Garcia (1986) [10]	Latency between 87 lowercase digraphs and space key and complex discrimination using Mahalanobis distance function	Individual's name and 1000 common words 10 times each	(N/A)	FAR 0.01% (N/A) FRR 50% (N/A)
Young and Hammon (1989) [11]	Plurality of features including: digraph latencies, time to enter selected number of keystrokes and common words using Euclidean distance	(N/A)	(N/A)	(N/A)
Joyce and Gupta (1990) [5]	Digraph latencies between reference strings using mean and standard deviation of latency distance vectors	Username, password, first name, last name 8 times each	33 users of varying ability	FAR 0.25% (2/810) FRR 16.36% (27/165)
Brown and Rogers (1993) [12]	Latencies and hold times using Euclidean distance and neural networks	Usernames, 15–16 character avg. ≈ 1,000 sequences tested	21 and V 25 users	FAR 4.2%–11.5% (N/A) FRR (N/A)
Obaidat and Macchiarolo (1993) [6]	Digraph latencies between reference strings using neural networks	15 character phrase 20 times each	6 users	97% overall accuracy
Obaidat and Sadoun (1997) [7]	Digraph latencies and key hold times using multiple machine learning algorithms	Username 225 times/day for 8 weeks	15 users	FAR 0% (N/A) FRR 0% (N/A)
Monrose and Rubin (1997) [1]	Latencies and durations with normalized Euclidean distance and weighted/nonweighted maximum probability	Passages of text over 7 weeks	(N/A)	Identification framework
Maisuria and Ong and Lai (1999) [13]	Digraph latencies with neural networks (multi-layer perceptron)	passwords 60 times over 3 periods	20 users	FAR ≈ 30% (N/A) FRR ≈ 15% (N/A)
Monrose, Weiter, and Wetzel (2001) [8]	Digraph latencies and key hold times, algorithm employed is unclear	8 character password	20 users	FAR % (N/A) FRR 45% (N/A)
Bergadano, Gunetti, and Picardi (2002) [9]	Trigraph duration using degree of disorder	683 character text 5 times	44 users	FAR 0.04% (1/10,000) FRR 4% (N/A)
Yu and Cho (2004) [14]	GA-SVM's and wrapper FSS on hold times and digraph intervals	6–10 character passwords 150–400 gen/user and 75 imp	21 users	FAR 0% (N/A) FRR 3.69% (N/A)
Bartlow and Cukic (2006) [2]	Random Forests on digraph latencies and hold times digraph latencies	usernames + 8 and 12 char passwords ≈ 9,000 sequences	41 users	FAR 2% (N/A) FRR 2% (N/A)
Sung and Cho (2006) [15]	GA-SVM's and wrapper FSS on hold times and digraph intervals	6–10 character passwords 150–400/user and 75 imposter	21 users	FAR 3.85% (N/A) FRR 13.10% (N/A)

performance is typically measured by conventional error measures such as False Accept Rate (FAR), False Reject Rate (FRR), and Equal Error Rate (EER). In terms of EER, many of the previously cited works achieve performance $\leq$5% (see Table 1). Naturally, FAR and FRR's can be tailored based on where one wishes to fall on a traditional Receiver Operating Characteristic (ROC) curve. It is important to note that the literature

has not firmly established whether the technology is sufficient for biometric systems operating in identification mode as the focus of past research is almost exclusively tailored to verification based systems. It is also important to note the trend of decreasing data requirements as earlier works required extremely long passages of text whereas most recent works require only usernames, passwords, or both. Related to this trend, keystroke dynamics need not be applied only at the time of login, which may lead to time-of-check-time-of-use vulnerabilities. Instead, they can be applied transparently throughout the span of a period of use. This feature can allow systems to continually check for the presence of insider threat where an authorized user may login to a system and subsequently allow an unauthorized user access. If a system does not require a continual verification environment, keystroke recognition is also very suitable for a ► challenge response type framework where the user is periodically authenticated.

Besides stand-alone biometric systems, keystroke recognition can be used as an augment to traditional username/password systems. This process is often called ► credential hardening or password hardening. Monrose et al. first proposed the idea [8] and Bartlow et al. also explored the concept [2]. Both works show how the addition of keystroke recognition to traditional authentication mechanisms can drastically reduce the penetration rate of these systems. Works of this nature may also bode well in online authentication environments such as banking and e-commerce websites which now commonly require secondary verification layers.

Either as a stand-alone biometric or an augment to a traditional username/password scheme, keystroke dynamics are arguably more cancelable or replaceable than physiological biometrics. The idea of cancellable biometrics touches on the fact that the threat of biometric compromise exists and is often realized. With fingerprint, face, iris, etc., it is often difficult to reissue a biometric authentication mechanism as fingers, faces, and irises are not easily removed and replaced in humans. In keystroke recognition however, the behavior which induces the biometric can be changed. In other words, if a user's keystroke recognition template is compromised, the data in which the template is based (i.e., password/passphrase) can simply be changed which will result in a new biometric template. For obvious reasons, this is seen as a very attractive feature of keystroke recognition.

Beyond the scope of academic research, many patents have been issued in the field including: Garcia (4,621,334 - 1986) [10], Young and Hammon (4,805,222 - 1989) [11], Brown and Rogers (5,557,686 - 1996), and Bender and Postley (7,206,938 - 2007). In addition to patents, there are many commercial offerings of keystroke recognition systems. Two popular systems are BioPassword ©(http://www.biopassword.com/) and iMagic Software ©(http://www.imagicsoftware.com). Systems such as these are attractive as the overhead of keystroke recognition in terms of hardware deployment and seamless integration into currently existing authentication systems is typically much less than that associated with physiological biometrics such as fingerprint, iris, and face.

Despite the maturity of the field over the last 30 years, there are still many challenges that are yet to be solved. Three main challenges are associated with the data required to train keystroke recognition systems. First, few works have formally set out to determine the amount of sequences required to sufficiently establish a typing signature ready for operational deployment. For a system to be deployable, it must have a realistic training requirement that the users are willing to incur. It seems that repeatedly typing a username and password combination 50 or more times would be unacceptable in the eyes of most users, yet five may be insufficient in terms of meeting established security goals. Second, as passwords need to be replaced or reissued, the problem of retraining needs to be addressed. Once again, these retraining requirements are yet to be firmly established. Third, the behavioral nature of this keystroke recognition requires a slightly more involved data collection process than what is typical in conventional physiological biometric systems. Most notably, one cannot simply compare genuine input of one user to genuine input of another user in order to establish an instance of imposter input as the data is often different for every user (i.e., usernames/passwords). As a result, most academic research will have users type the credentials or data associated with other users to arrive at imposter sequences for training. Clearly this is not feasible in operational systems as passwords are frequently reset. Therefore, the issue of automatic generation of imposter data is an area that needs to be explored.

Summary

Keystroke recognition is a behavioral biometric which authenticates an individual not on the basis of what is typed but the nature of how it is typed. A large base of research has accumulated in the field over the last 30 years establishing its potential both as a stand-alone biometric and an augment to traditional username/password authentication schemes. Due to its transparent nature, low cost of deployment, and seamless fit into currently existing commercial and governmental applications, it is an excellent candidate for increasing the security of authentication systems.

Related Entries

▶ Biometric Encryption
▶ Cancelable Biometrics
▶ Verification

References

1. Monrose, F., Rubin, A.D.: Authentication via Keystroke Dynamics. In: ACM Conference on Computer and Communications Security, pp. 48–56 (1997)
2. Bartlow, N., Cukic, B.: Evaluating the Reliability of Credential Hardening through Keystroke Dynamics. In: ISSRE, IEEE Computer Society, Washington, DC, USA, pp. 117–126 (2006)
3. Mueller, S.: Upgrading and Repairing PCs, 15th edn. QUE, Indianapolis, IN (2004)
4. Gaines, R., Lisowksi, W., Press, W., Shapiro, S.: Authentication by keystroke timing: Some preliminary results. Rand Report R-256-NSF, The Rand Corporation, Santa Monica, CA (1980)
5. Joyce, R., Gupta, G.: Identity authentication based on keystroke latencies. Commun. ACM **33**(2), 168–176 (1990)
6. Obaidat, M.S., Macchairolo, D.T.: An on-line neural network system for computer access security. IEEE Trans. Industrial Electronics **40**(2), 235–241 (1993)
7. Obaidat, M.S., Sadoun, B.: Verification of computer users using keystroke dynamics. IEEE Trans. Syst. Man Cybern. **27**(2), 261–269 (1997)
8. Monrose, F., Reiter, M.K., Wetzel, S.: Password hardening based on keystroke dynamics. Int. J. Inf. Sec. **1**(2), 69–83 (2002)
9. Bergadano, F., Gunetti, D., Picardi, C.: User authentication through keystroke dynamics. ACM Trans. Inf. Syst. Secur. **5**(4), 367–397 (2002)
10. Garcia, J.: Personal identification apparatus. Patent 4,621,334, US Patent and Trademark Office, Washington, DC (1986)
11. Young, J., Hammon, R.: Method and apparatus for verifying an individuals identity. Patent 4,805,222, US Patent and Trademark Office, Washington, DC (1989)
12. Brown, M., Rogers, S.J.: User identification via keystroke characteristics of typed names using neural networks. Int. J. Man Mach. Stud. **39**(6), 999–1014 (1993). DOI http://dx.doi.org/10.1006/imms.1993.1092
13. Maisuria, L.K., Ong, C.S., Lai, W.K.: A comparison of artificial neural networks and cluster analysis for typing biometrics authentication. In: International Joint Conference on Neural Networks (IJCNN), vol. 5, pp. 3295–3299 (1999)
14. Yu, E., Cho, S.: Keystroke dynamics identity verification - its problems and practical solutions. Comput. Secur. **23**(5), 428–440 (2004)
15. Sung, K.S., Cho, S.: GA SVM wrapper ensemble for keystroke dynamics authentication. In: ICB, Springer-Berlin-Hiedelberg, pp. 654–660 (2006)

Kinematic Body Model

Virtual skeleton structure comprising a fixed number of joints with specified angular degrees-of-freedom. The values assigned to these joint angles define the 3D pose of the body.

▶ Markerless 3D Human Motion Capture from Images

Kinematics

The description of object motion over time, generally expressed in terms of position, velocity, and acceleration.

▶ Human Detection and Tracking

Knowledge-based Gait Recognition

▶ Gait Recognition, Model-Based

Known Traveler

▶ Registered Traveler

L

L2 norm

L2 norm is a standard method to compute the length of a vector in Euclidean space. Given $x = [x_1\ x_2 \ldots x_n]^T$, L2 norm of x is defined as the square root of the sum of the squares of the values in each dimension.

► Iris Super-Resolution

Lambertian Law

Lamberts cosine law states that the reflected or transmitted luminous intensity in any direction from an element of a perfectly diffusing surface varies as the cosine of the angle between that direction and the normal vector of the surface.

► Face Recognition, Near-infrared
► Heterogeneous Face Biometrics

Lambertian Surface

Lambertian surface is a technique used to light particular surfaces of virtual objects within a scene, which causes all closed polygons to reflect light equally in all directions. This means that the surface brightness to an observer is the same regardless of the observer's angle of view.

► Face Sample Quality
► Illumination Compensation

Large Scale Biometric Database

Based on their size all biometric datasets are classified into three (some sources indicate four) categories: small size, medium size, large scale, and very large scale data bases. The size of datasets is determined by the number of participating users. A small size database can contain biometric data of up to 1,000 users. A medium size database accounts for 10,000–100,000 users. A large (and very large) scale datasets include biometric data of more than 1,000,000 users. Since each user can be represented by two or more classes in a database, some references indicate the size of datasets in classes rather than in users.

► Face Databases and Evaluation
► Iris Sample Synthesis
► Large Scale System Design

Large Scale Biometric System Design

► Large Scale System Design

Large Scale System Design

Chin-Hung Teng[1], Wen-Hsing Hsu[2]
[1]Yuan Ze University, Taiwan, ROC
[2]National Tsing Hua University, Taiwan, ROC

Synonym

Large scale biometric system design

© 2009 Springer Science+Business Media, LLC

Definition

A large scale biometric system is a system involving the authentication of a huge number of users via the biometric features. A Large Scale ► Biometric Database is generally designed for civilian applications and is not merely the increased size of database compared to the personal use system. In the case of a large scale system, there is greater emphasis on the issues of system reliability and flexibility. To meet the requirements of public use, the system must have a high enrollment rate and adapt to environmental variations. Security is another issue in designing such a system due to the large number of users enrolled. A one-to-many matching may sometimes lead to system breakdown, thus other authentication policies may be applied to complement the one-to-many matching to enhance system reliability, such as cross validation of multiple biometric features.

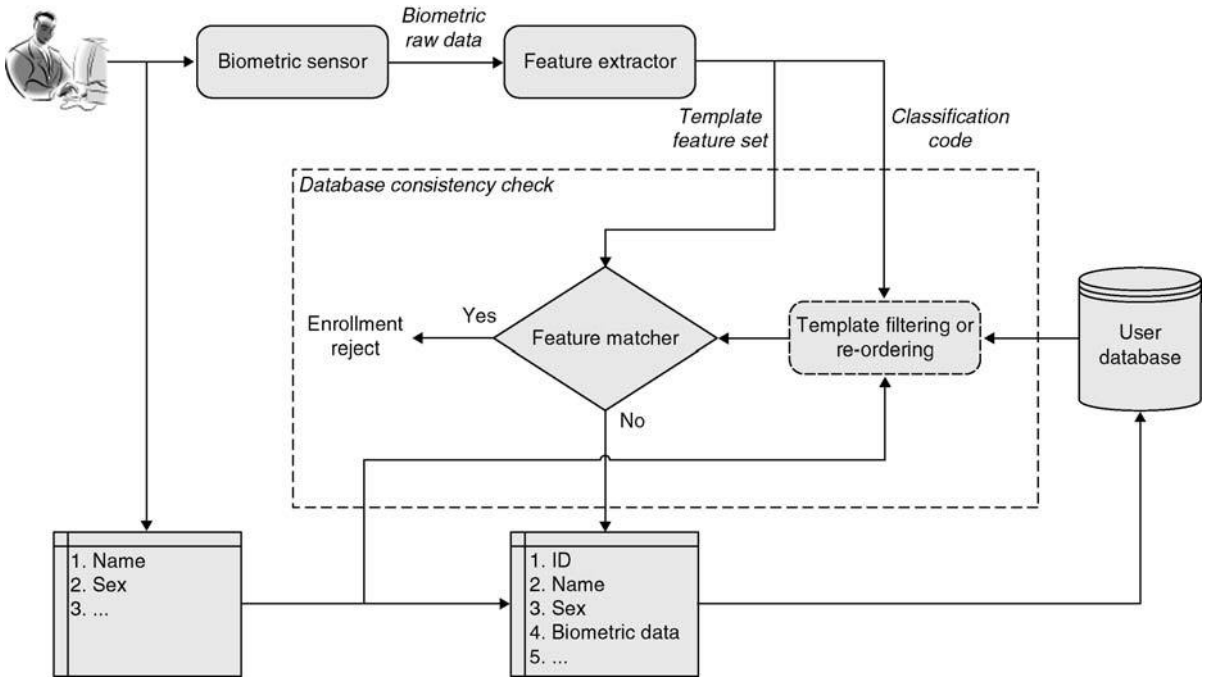
Introduction

There is a long history of fingerprint verification/identification in law enforcement, a large scale biometric system which is primarily used for determining the identity of a suspect or a dead person. In the early stage, fingerprint verification is achieved by an expert who must visually match the ► minutiae of different fingerprints, which is undoubtedly a tedious and time consuming task. In recent years, due to advances in computer recognition techniques and the extensive development of biometric sensor devices, the identity of an individual can be authenticated through biometric features fully automatically by computers. A biometric system has the distinct advantages of high security and convenience, thus theoretically it can find many applications where the users need to prove the claimed identity to access the required services. However, due to the issue of human rights (it is unfortunate that a fingerprint is always associated with crime), many people are unwilling to provide their biometric features in public systems, thus most biometric applications are at present restricted to personal use such as notebook or mobile phone login. After 9/11 attacks, many people have realized the importance of security and this has pushed the government to initiate many large scale biometric systems to enhance the security for access in the government's facilities and information systems. In fact, a biometric system is not merely used for preventing terrorist attacks. Many government provision such as welfare disbursement, driver license application, and voter registration can benefit from the construction of a nationwide biometric system. For example, many airports have installed a biometric authentication system to expedite the procedure of visa and passport examination. The United States government has launched a Personal Identity Verification (PIV) program [1] to manage the authentication of federal employees and contractors for access to federal facilities and information systems. It is obvious that in the future more and more large scale biometric systems will be developed. This essay outlines some issues that should be considered in designing a large scale biometric system and lists some of the applications of a system.

Operation of a Large Scale Biometric System

Typically, there are two authentication modes for a biometric system: verification (one-to-one matching) and identification (one-to-many matching). Verification is a procedure for comparing a biometric feature set against a template with claimed identity. For a large scale system, the enrolled template can be stored in a centralized database, a set of distributed databases (sometimes at distant places), or a user carried medium such as an IC card, depending on the requirements of applications. Different arrangements have their own advantages and disadvantages. For instance, in the PIV program, the biometric templates are distributed to the user's PIV card, thus saving effort in managing the huge number of records of biometric data. However, the data stored in the card must be carefully protected so as to prevent an impostor replacing or stealing the biometric data to crack the system. Usually, the protection mechanism is achieved by a Public Key Infrastructure (PKI).

Identification is a procedure for comparing a biometric feature set against all templates in the database to determine the correct identity. Generally, if a biometric system cannot achieve a sufficient low False Matching Rate (FMR), it is not recommended to use the one-to-many matching to guard access to some services or resources since it will greatly reduce the security level and result in a long matching time. However,



Large Scale System Design. **Figure 1** The enrollment module of a large scale biometric system.

identification is at times necessary for some applications such as detecting multiple enrollments in a system. If an application requires identification in a large scale biometric system, some strategies must be applied to reduce the searching size so as to enhance security as well as to expedite the matching process. An enrollment process for a large scale system with the functionality of detecting duplicate enrollments is presented in Fig. 1. To reduce the number of matchings, some reliable information about the applicant can be used to filter out incorrect biometric templates. For example, it is possible to cut out a large portion of templates to be matched by the sex of the applicant. If other information such as the applicant's eye color can also be used to reliably classify the users, the search size can be further reduced. Normally, a biometric classification mechanism can also be included in the system to accelerate the identification process. For instance, a fingerprint can be categorized on the basis of the ridge pattern into arch, right loop, left loop, whorl, etc. Typically, if a classification system cannot achieve a sufficiently high level of accuracy to effectively cut out the templates to be matched, classification is just a re-ordering of the templates that increases the probability of matching the biometric templates from the same individual as soon as possible. This implies that a classification system can efficiently detect duplicate enrollments if the applicant has enrolled in this system previously, but for a new applicant, re-ordering of matching templates has no benefit to the system since the biometric template of the applicant is not in the database.

Some Issues for Designing a Large Scale Biometric System

In general, designing a large scale biometric system is different from personal use systems. Some issues should be carefully considered.

1. Cross-sensor problem and the development of standards: For a personal use biometric system, the biometric features for enrollment and verification are usually acquired from the same biometric sensor. This obviously rules out the problem of cross-sensor matching which sometimes leads to performance degradation due to sensor discrepancy. However, cross-sensor matching is quite common in a large scale system where the enrollment and verification sensors are typically different. Sensor discrepancy is due to several factors such as the manufacturing process, sensors developed by different vendors, and even the different sensor design methodologies

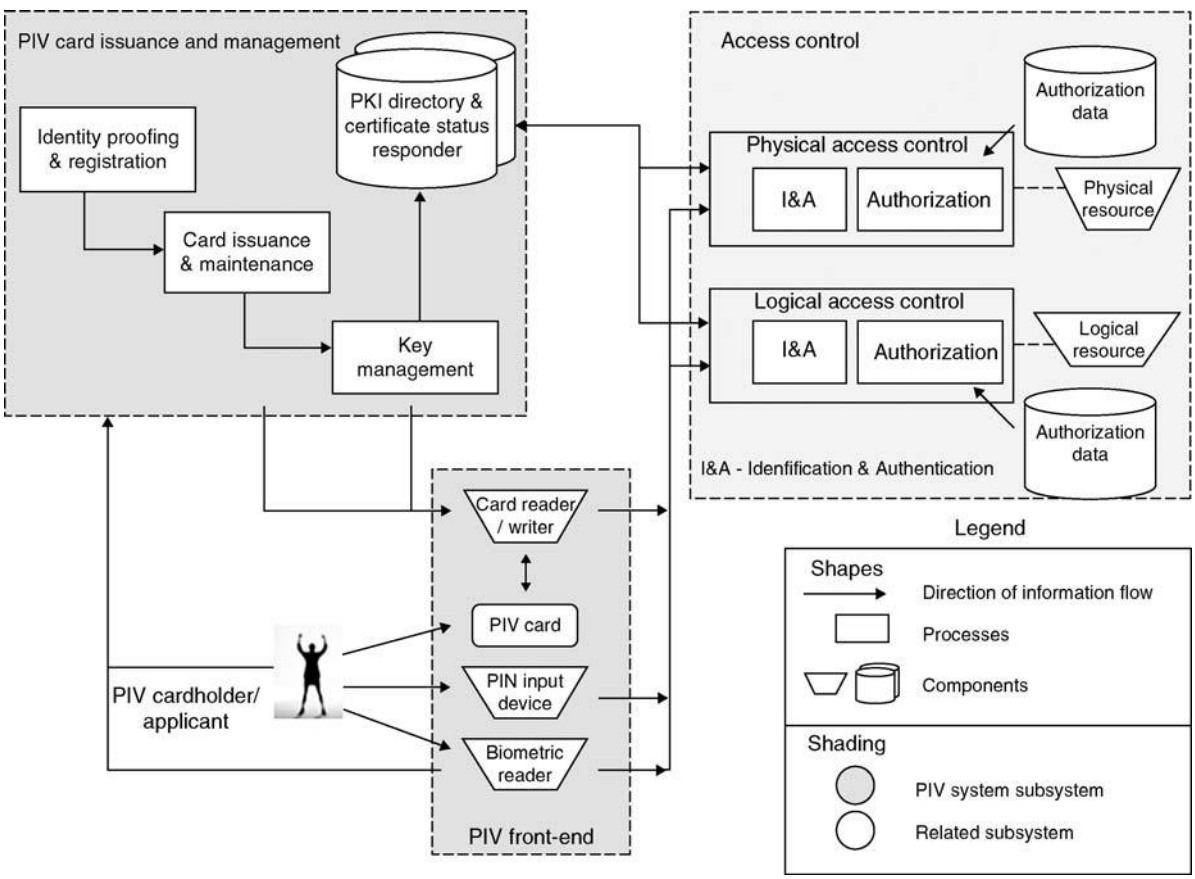
such as optical-based and chip-based fingerprint sensors. One way of solving the problem of cross-sensor matching is to establish a standard to regulate the quality of captured biometric data such as the FBI's fingerprint image quality specification [2]. Biometric standards, however, are not restricted to the area of defining the quality of biometric data. The standards for the format of biometric template and the ►application programming interface (API) have also been developed. Not only, fingerprint image quality standard has been established, but also biometric API has been developed such as BioAPI [3]. In fact, many standards for the security industry, such as the interface of access control, audio verification, and control panels, have also been developed. The Open Systems Integration and Performance Standards (OSIPS) of the Security Industry Association (SIA) [4] is a well-known program for the development of security standards. Under the OSIPS, a series of standards have been published to meet the requirements of emerging IT products and services in the security industry. These standards allow different manufacturers to cooperate for designing very large scale biometric systems and create more business benefit from security applications. It is likely that in the future only those biometric products that comply with these standards may be accepted by the market.

2. Poor biometric quality in the case of specific users: It has been known that the quality of biometric features of some specific users is very poor, especially those engaged in particular occupations. For instance, the fingerprints of some porters are generally worn to such an extent that it is quite difficult to recognize their fingerprints. Because a large scale biometric system is generally designed for public use, it cannot reject those users with poor biometric quality or those with biometric deficiency. In addition to non-technique complement schemes, a biometric system should further improve their performance, including biometric feature enhancement as well as matching algorithm, to accommodate those users with poor biometric features. This system should generate high quality biometric data or templates for a very large proportion of the user population.
3. Environmental variations: Environmental variations such as illumination or temperature variations sometimes lead to performance degradation of a biometric system because of unequal sensor conditions of the captured biometric templates to be matched. For instance, face recognition is typically quite sensitive to ambient illumination variations. For a personal use biometric system, biometric sensors are generally set up in an indoor and controllable situation so as to minimize environmental variations. However, for a public use biometric system, sensors may be set up outdoors, thus any biometric quality degradation that may be due to environmental variations should be taken into account when designing a large scale biometric system. A large scale system is expected to generate high quality biometric data across the full range of environmental variations for the intended applications.
4. Performance degradation for one-to-many matching: As outlined previously, one-to-many matching may have performance degradation. For example, if a biometric system has 0.0001 false matching rate (FMR) and there are 10,000 biometric templates in the database, then a one-to-many matching against this database will always produce a successful matching with a statistical probability of 1. Thus, when applying a one-to-many matching in a large scale system, system designer must carefully compute the resulting error rate and check whether it satisfies the demand of the intended application. If a one-to-many matching is inevitable for a large scale application, several mechanisms can be included in the system to enhance the security level. For instance, it is possible to reduce the probability for false alarm matching by cross validation of multiple biometric features. Multiple biometric features may be obtained from the same biometric type such as fingerprints of different fingers, or from different types such as face, voice, fingerprint, and hand geometry etc. Regardless of which biometric features are used for cross validation, the resulting security level should be carefully calculated to ensure that it meets the security requirements. Beside a degraded error rate for one-to-many matching, the increased matching time may at times become a design issue for a large scale system. If cost is not a critical factor, the matching time can be greatly reduced by distributing the matching task into several matching machines. An efficient classification scheme can also be used to reduce the matching time. If a classification algorithm can effectively filter out the unnecessary templates to be matched, it can also improve the error rate of one-to-many matching.

The PIV Program of the United States Government

The PIV program of the United States government is a large scale biometric authentication project intended to control the federal employees and contractors for access to federal facilities and information systems. Under this program, every applicant is issued a PIV card on which two types of biometric information are recorded: one is the photograph of the applicant which is printed on the card and the other is the applicant's fingerprint template which is stored in the card memory. The photograph of the applicant is used for visual-based authentication which serves as the lowest level of identity assurance. This visual-based authentication is necessary in situations where electronic biometric authentication is not workable, e.g., the PIV card readers cannot be installed on the site. In other situations, authentication is automatically achieved by comparing the cardholder's fingerprint against the template stored in the card. The system notional model of the PIV program and its conceptual operation are summarized in Fig. 2.

The PIV program is open to biometric manufacturers worldwide, thus any biometric vendor whose techniques comply with the requirements of the PIV can join this program. For this purpose, the Department of Commerce and the National Institute of Standard and Technology (NIST) have published a series of standards to regulate the biometric techniques. The NIST has also designed a performance evaluation test to cross verify the fingerprint verification techniques developed by different manufacturers [5]. In this test, the entire process of fingerprint verification is divided into two phases: feature extraction and template matching. The goal of this test is to ensure the reliability of extracted fingerprint features and the capability of fingerprint matching algorithm. The false matching



Large Scale System Design. **Figure 2** The system notional model of the PIV program. (Reprinted from Personal Identity Verification (PIV) of Federal Employee and Contractors, Federal Information Processing Standards Publication).

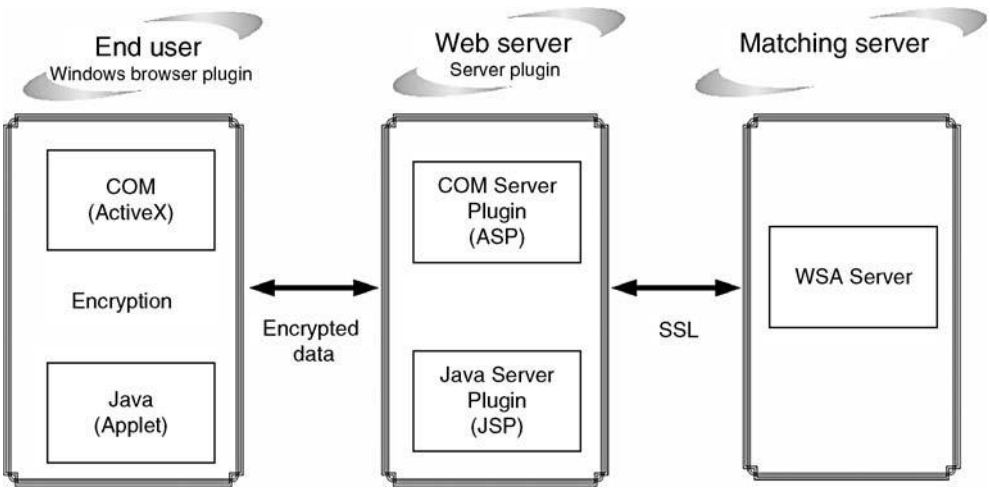
rate (FMR) and the false non-matching rate (FNMR) for pairs of feature extraction module and fingerprint matching algorithm from different vendors are evaluated to check whether they can achieve the desired security level. This test does not only examine the compatibility of biometric techniques developed by different vendors but also checks the fingerprint recognition accuracy of the vendors. Only those biometric manufacturers who pass the test can serve as the biometric techniques providers for the PIV program.

Web Service Authentication (WSA)

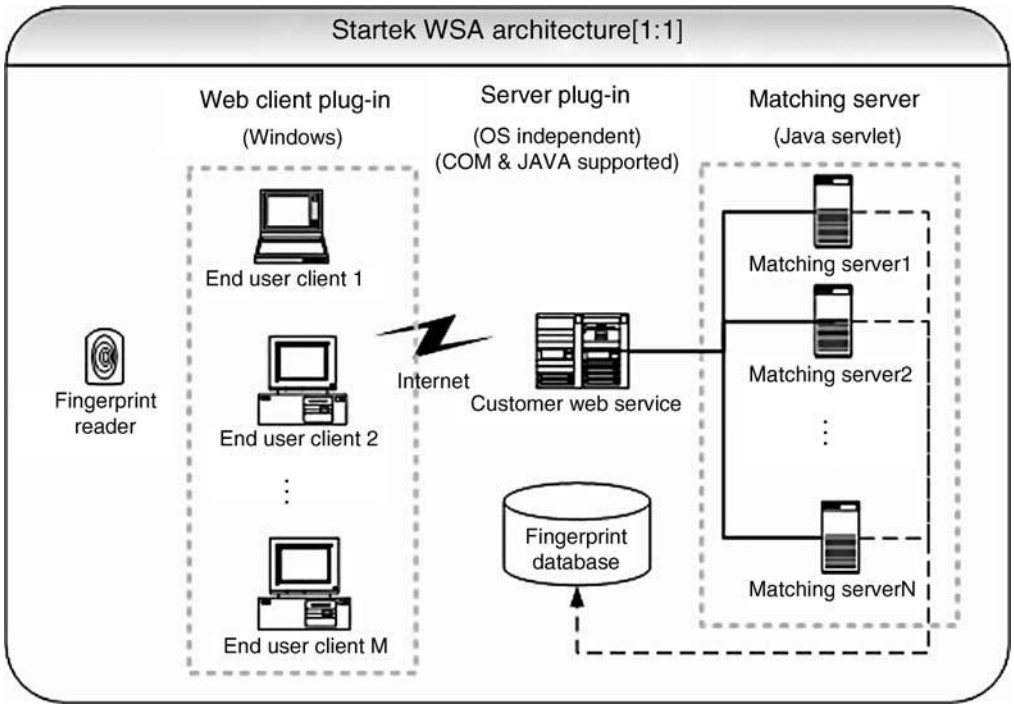
With advances in internet technology, a biometric authentication system can also be applied to the browser to replace the traditional password-based login system (i.e., a web service biometric authentication), leading to a more convenient and secured biometric solution on the internet. This web service authentication is not limited to biometric applications of personal use but can be extended to a large scale biometric system such as a web-based time-attendance system for a global company with a number of branches around the world. Each employee can register his/her attendance time via this web service authentication system even when he/she is in a remote branch. Typically, there are three components of a web service authentication system as illustrated in Fig. 3. The terminal of End User is usually a browser with some plugins (ActiveX in IE or Java applet for other platforms) in charge of connecting the biometric sensor and sending the extracted biometric template to the Web Server. The Web Server is responsible for managing the biometric templates and sending the templates to be matched to the Matching Server. The Matching Server deals with the requests from different Web Servers (which may be devised for various functions and applications) and respond with the authentication results. All the communication data between servers and browsers are transmitted via a secure channel which is typically achieved by a data encryption scheme. In this framework, each component has its distinct functionality and role, thus the overall system is easy to manage and maintain. In addition, since the plugins can be installed directly by the internet, setting up a web service authentication system is fairly simple. A biometric reader installed on a computer with web-access functionality is sufficient for completing the system. A more general framework is depicted in Fig. 4. The matching servers can be extended to a matching array thus significantly enhancing the matching efficiency. In short, this web service authentication scheme is easy to install, easy to manage and maintain, platform independant, flexible function extension, and easy to integrate to other systems, thus it is likely to be widely used in the area of biometric recognition on the internet.

Summary

This essay presents a brief description of the operation of a large scale biometric system and highlights several



Large Scale System Design. **Figure 3** The three components for a web service authentication scheme (Data Source: Startek).



Large Scale System Design. **Figure 4** A more general architecture for web service authentication (Data Source: Startek).

L

issues in designing such a system. It discusses a large biometric project, the PIV program of the United States government, and also examines a framework for biometric application on the internet. Besides, it enumerates the advantages of such a biometric network application. The authors of this contribution aims to provide a simple view to those who seek an initial understanding of the design of a large scale biometric system.

Related Entries

- ▶ Authentication
- ▶ Automated Fingerprint Identification System
- ▶ Biometric
- ▶ Enrollment
- ▶ Identification
- ▶ Verification

References

1. Personal Identity Verification; information available at: http://csrc.nist.gov/piv-program/
2. Test Procedures for Verifying IAFIS Image Quality Requirements for Fingerprint Scanners and Printers, MITRE Corporation Technical Report, MTR-050000016, April 2005. Document available at: http://www.mitre.org/tech/mtf
3. BioAPI Consortium; information available at: http://www.bioapi.org/
4. http://www.siaonline.org
5. http://fingerprint.nist.gov/minexII/index.html
6. Personal Identity Verification (PIV) of Federal Employees and Contractors, Federal Information Processing Standards Publication
7. http://www.startek-eng.com/EN/prod_WSA20Solution.html

Large-Scale Evaluation

Large-scale evaluation is the evaluation that involves testing on significant amounts of data, that is, Large Scale Biometric Databases. It normally provides results using the statistical measurements, such as average FAR, FRR and/or ROC, and CMC curves.

- ▶ Face Databases and Evaluation

Latent Fingerprint

A fingerprint left on an object by touching it. Example objects are glasses, doors, and tables.

- ▶ Biometric Identification
- ▶ Fingerprint Features
- ▶ Latent Fingerprint Experts
- ▶ Law Enforcement Agency
- ▶ Liveness Detection: Fingerprint
- ▶ Security and Liveness, Overview
- ▶ Universal Latent Workstation

Latent Fingerprint Experts

Thomas A. Busey, Bethany L. Schneider
Psychological and Brain Sciences; Program in Cognitive Science, Indiana University, Bloomington Indiana, IN, USA

Synonym

Latent fingerprint recognition

Definitions

Cognitive processing is the term given to mental effort directed toward a particular problem. Cognitive science is an umbrella term given to all disciplines that focus on intelligent systems; research psychologists traditionally focus on human performance. The field of cognitive science includes mathematicians, computer scientists, research psychologists, biologists, and philosophers. Cognitive processing is closely linked to perceptual processing and decision making, both of which are involved in latent print examinations. As part of the science, researchers typically collect data from experts and novices to document how and when expertise develops.

▶ Latent fingerprint examiners are practitioners who are trained to individualize or exclude latent prints and prints from known sources (e.g., 10-print cards). These practitioners often work with automated databases such as IAFIS, which provide candidate prints from known sources.

Introduction

Engineers attempt to solve a biometric problem by isolating features or dimensions that they believe are diagnostic, or use machine learning procedures to identify a feature set that might be useful. Cognitive scientists take the opposite approach. They use testing procedures designed to infer the brain processes that underlie performance in human experts. Under the assumption that humans have the most flexible information processing system and can use different levels of information, this reverse engineering approach holds the promise of improved quantitative analyses of fingerprints. This entry summarizes the work that has characterized performance in latent print examiners, and describes how cognitive and vision scientists design experiments to reveal the mechanisms underling human latent print identification.

The study of expertise in latent print examiners is a relatively new field, and only a few group has published on the topic. However, there is a great deal of research in related fields, and this research is described where it applies to latent print examinations.

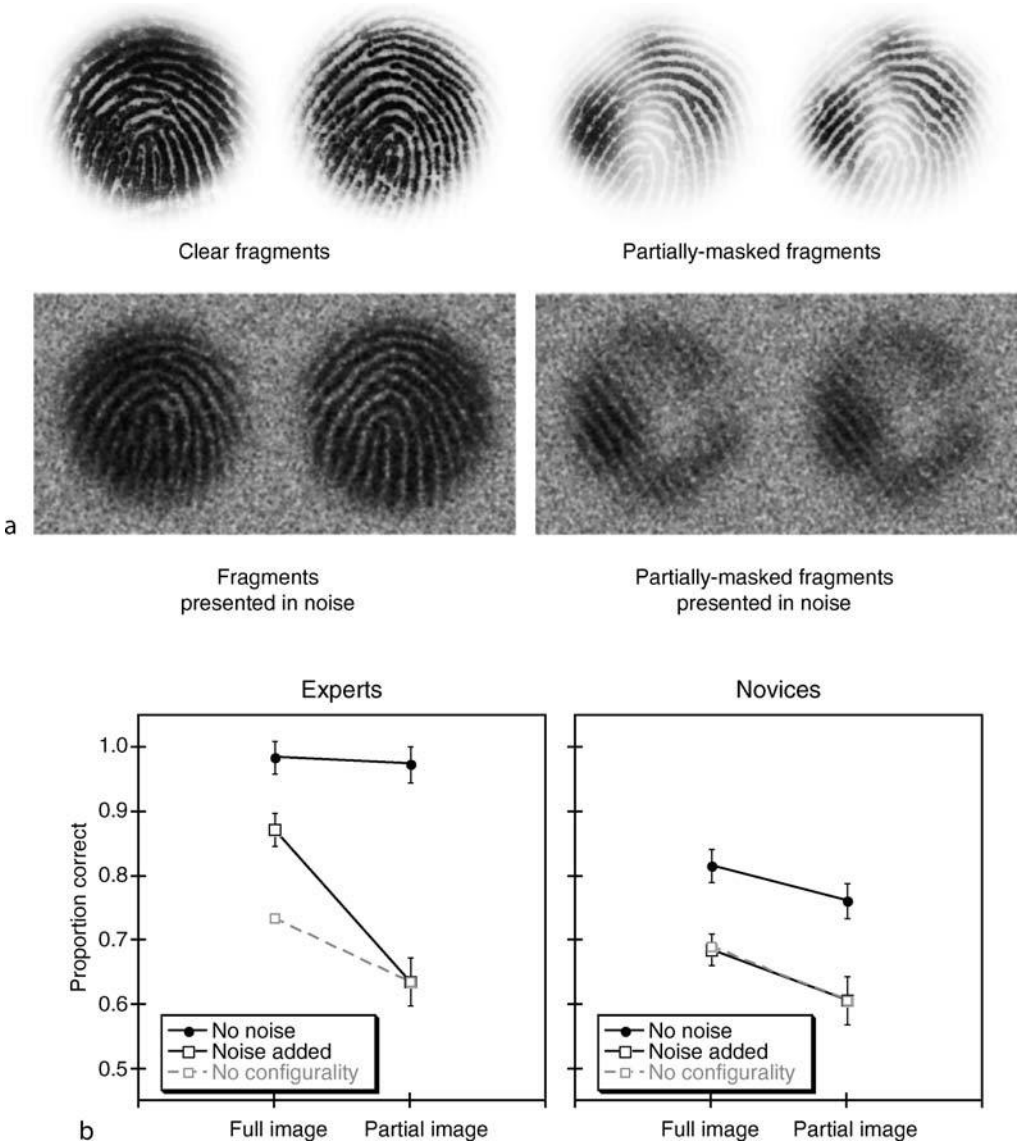
Empirical Evidence

Research on ▶ perceptual expertise in human experts proceeds much in the same way that an engineer would evaluate the performance of a biometric system. Researchers generate candidate mechanisms that describe how an expert would accomplish a particular task. These candidate hypotheses are based on an analysis of the information available, along with known perceptual and memory constraints of humans. The "system" (in this case the human) is tested with a recognition or memory task and performance measures such as false match rate and false non-match rate can be computed. Because the difficulty of a particular task depends on the choice of materials, researchers often compare performance from human experts against those of human novices. Busey and Vanderkolk [1] tested experts and novices in a fingerprint fragment

matching task that was one of the first major studies of expertise in latent print examiners. Described below are the details of this study, which not only illustrates how experts differ from novices but provides an illustration of how research questions are developed and answered in Cognitive Science.

Stimuli such as those shown in Fig. 1 were presented briefly to expert latent print examiners and novices. A single print would be shown for 1/5th of a second, followed by a pattern mask for either 200 ms or 5 s. Then two prints would be shown, one of which exactly matched the studied print fragment. The test prints could either be whole or partially masked to simulate a latent print, and could be presented with or without noise (which simulates the fact that some latent prints are recovered on textured or marked surfaces which adds visual noise).

The data in Panel B of Fig. 1 illustrates several factors. First, experts perform better than novices, a difference that is even more pronounced at longer



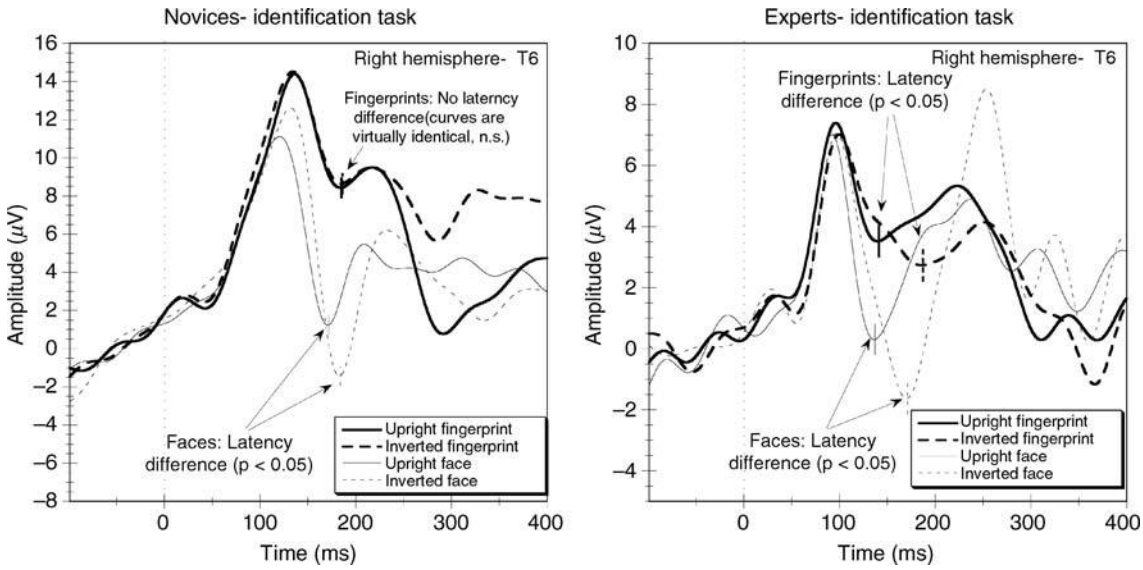
Latent Fingerprint Experts. Figure 1 Panel A: Stimuli used to test novices and latent print examiners [1]. Panel B: Empirical data demonstrating improved performance overall for experts, better performance in noise, and evidence for configural processing (see text for details).

delays (not shown). This demonstrates that experts may have better visual memories or the ability to re-code visual information into verbal descriptions which survive for longer intervals. Second, the structure of the design allows an investigation of ▶ configural processing. The partially masked prints have exactly half of the information in the full prints. Performance from the partially masked prints can therefore be used to predict what performance should be in the full print condition, using a model called *probability summation*. The dashed lines show the prediction of the probability summation model, and demonstrate that experts exceed that prediction for the prints in noise. This illustrates that experts are gathering more information from the second half of the print once it is added to the first half than one would have expected based on their performance on the first half. Another way to view this is that for experts, the whole is greater than the sum of the parts. One interpretation of these results is that experts use the information from one half to make better use of the information from the second half when both are present.

Converging evidence for configural processing was found using brain recording in a second experiment. When visual stimuli are presented, neurons in the brain fire and give off electrical activity than can be recorded by placing electrodes on the surface of the scalp. This form of electroencephalography (EEG) allows researchers to monitor the ongoing brain activity that is elicited by a visual stimulus. While this technique has only coarse spatial resolution due to the spreading nature of electrical charges, it has excellent temporal resolution, on the order of millisecond accuracy.

Researchers using this particular technology have noted that stimuli that are known to be processed using configural or holistic mechanisms such as faces and similar trained stimuli provide a signature of this configural processing [2]. Voltage recorded from the head and averaged over lots of trials provides the event-related potential (ERP), and faces produce a very distinctive feature over the left and right parietal regions of the brain. This feature is a downward-trending component that has an inflection at about 170 ms after stimulus onset. Figure 2 illustrates this feature, which has been termed the N170. When the stimulus is inverted, which has the effect of reducing or eliminating configural processing, the N170 is delayed and slightly more negative-going.

This signature of configural processing provides a means to test whether experts process fingerprints configurally. Experts and novices were shown upright and inverted faces and fingerprints. As expected, both experts and novices show differences between upright and inverted faces. However, only experts showed a similar pattern for fingerprints as they showed for faces: The N170 for inverted fingerprints was significantly delayed relative to upright fingerprints, but only in experts. The dark curves in Fig. 2 illustrate this effect. These



Latent Fingerprint Experts. **Figure 2** Electrophysiological data from Novices and Experts with upright and inverted fingerprints and faces [1]. Light curves come from faces, while dark curves are from fingerprints (see text for details).

findings are important because not only do they illustrate demonstrably different patterns of brain activity in experts and novices, but the time course of the differences are consistent with processing that happens relatively early in visual processing. Thus these examiners are experiencing relatively low-level changes in their visual system that improve the quality of the information and the way they interpret this perceptual information.

In summary, the behavioral and electrophysiological evidence from latent print examiners supports the view that experts have better recognition overall for fingerprints, they have better visual memories for fingerprint information, and they process fingerprint information in qualitatively different ways using configural processing mechanisms.

Perceptual Expertise

While little research has focused specifically on latent print examiners and the changes that develop as a result of their expertise, candidate mechanisms that have been previously discovered by cognitive scientists using related materials can be extended. For example, the idea of configural (using relational information between parts) and/or holistic (obligatory processing of all the parts of an object) processing has become a consistent theme throughout the literature and many researchers argue that it is a signature of expertise [3]. Specifically, researchers studying perceptual expertise have developed paradigms that test for and illustrate a shift from a feature-based system of object recognition (seeing individual parts of an objects) to the use of holistic and/or configural mechanisms [3]. These effects are often illustrated in behavioral tasks that train subjects on a specific stimulus type and then test these subjects on either the studied or transformed configuration or isolated parts. Post-training performance is often compared with either their pre-training performance or with novices (those who receive no training). The underlying theme that results from these research paradigms is that experts develop a holistic system which causes them to be more sensitive to configurations and be unable to ignore distractor parts of the stimulus.

Apart from establishing configural and/or holistic mechanisms, another key issue in expertise studies is showing how experience with a domain causes a reorganization of the visual recognition hierarchy away from the basic level and to the subordinate level. In general terms, subjects more readily identify items based on the basic level category membership (e.g., bird, table) rather than their subordinate membership (e.g., robin, coffee table) [4]. This hierarchy is structured to reflect the prominent use of basic-level information over the subordinate level information. However, a series of experiments has shown that the development of expertise results in enhanced subordinate level identification [5]. It has been proposed that (1) expertise causes a shift in the hierarchy to the subordinate level rather than the basic level, (2) experts make identifications based on this subordinate level information, and (3) their expertise allows them to be equally proficient in making identification on the subordinate and basic level. The proficiency in which experts use this subordinate level information has been reliably replicated and has been argued to be a signature of expertise.

Other studies have researched differences between experts and novices in terms of how expertise impacts ▶ visual memory, the ability to use verbal redescriptions, and attention to particular features. For example, previous studies on expertise have implicated enhanced visual memory for expertise items, and showed that chess masters were able to accurately reproduce VALID board configurations after viewing them for only 5 s [6]. This is arguably due to their extensive knowledge of specific patterns that results with expertise in the domain. Such an idea can be applied to latent print examiners and has also been reported for experts in other domains such as bridge players [7], music students [8], and electronics technicians [9]. In addition, this idea can be extended into the category learning literature by a finding that shows increases in memory sensitivity account for the ability to learn to uniquely identify similar objects [10]. This enhanced memory ability could also be linked to an enhanced ability to fixate on features that are the most informative for future identification, an idea that is also supported in the category learning literature [11].

In addition to visual memory, research in the perceptual categorization literature argue that experts develop a more robust storage, such as implicit verbal redescriptions, in the process of specializing in a category. Specifically, experts appear to garner more verbal knowledge about a domain but make categorizations without explicit deliberations [12].

Research with radiologists suggests that expertise may alter what types of perceptual information are

allowed for consideration. Specifically, experts arguably attenuate to specific task relevant dimensions [13]. More generally, this idea has also been founded in the category learning literature by showing that category learning includes learning how to optimally allocate attention to those features relative to the category and/or task and discard unrelated features [11].

Decision Making and Decision Biases

Expertise brings special abilities, but it also can lead to special vulnerabilities. Several studies conducted by Itiel Dror and co-workers illustrate the role that context plays in decisions about fingerprint individualizations. This can bring about ▶ contextual biases. The difficulty with latent print examinations is that the judgment that two prints come from the same source is essentially based on similarity. Even if one source (say a fingerprint) maintains a persistent structure over time, the way that print is laid down can greatly affect its appearance. Thus no two impressions of a single source will look identical. Clear prints might look very similar, but in the end individualization essentially comes down to a judgment that two prints look more similar than any close non-match that the examiner has seen. Such a task has three possible decisions (individualization, exclusion, or insufficient detail to make a determination). A particular pair of prints will produce some amount of evidence for each of these decisions, but whether the evidence exceeds some internal threshold depends on the individual examiner. Dror et al. explored the possibility that the details of the case (the context) might affect the decision process.

In the first study [14], non-experts where shown pairs of fingerprints and given additional (fictitious) details about the case. When the prints were shown with a highly emotional context such as an accident scene picture or a murder victim, the stimulus affected the decision made by subjects. Subjects were more likely to report a matching fingerprint pair when the context was emotional. This suggests that contextual information beyond the particular fingerprint perceptual information plays a role in latent print examinations, at least with novices.

To extend this to experts, two additional studies used covert measures to assess the role of context with examiners during their normal workflow [15, 16]. To highlight the importance of this work, consider the task of an examiner. He or she must evaluate the perceptual evidence and decide whether there is sufficient evidence to make a decision. What constitutes "sufficient" is of course of primary importance. Research psychologists refer to this task as a criterion-based (as opposed to a criterion free) judgment, since the decision outcome is based in part on the criterion that the examiner establishes. If the examiner allows additional details about the case that are irrelevant to the particular identification at hand to influence their criterion, they reduce value and independence of the latent print examination.

In this particular set of studies, latent prints from closed casework were given again to the same experts under the guise of a new case. These prints had previously been matched or excluded by the examiners. Dror and his colleagues found that 8 out of 11 experts made a decision that was inconsistent with their previous decisions on the identical pairs of prints. Most of the switched decisions occurred with difficult prints that were previously judged as identifications, although some of these easy identifications also had changed answers. The details of these experiments are complex, and the reader is referred to the original sources for full details [17], but the implications are clear: context can play a role in the decision that an examiner makes, and care must be taken not to allow external influence to affect the perceptual judgment.

Summary

Research on latent print examiners has demonstrated increased recognition of latent prints and the flexibility to rely on different levels of print information. These differences are supported by superior visual memory and different styles of process. The EEG data suggest that training and experience are causing changes in relatively early and low-level areas of the visual system that improve the quality of the perceptual information. A host of related studies suggest that experts may learn to re-code visual information into verbal descriptions, and learn to attend to the most relevant and diagnostic regions of a print. However, with these increased abilities may come increased vulnerabilities such as contextual biases that may affect the interpretation of a fingerprint pair.

Related Entries

- ▶ Fingerprint Classification
- ▶ Fingerprint Matching, Manual
- ▶ Fingerprint Recognition, Overview
- ▶ Law Enforcement
- ▶ Multiple Experts
- ▶ Psychology of Gait and Action Recognition

References

1. Busey, T.A., Vanderkolk, J.R.: Behavioral and electrophysiological evidence for configural processing in fingerprint experts. Vis. Res. **45**, 431–448 (2005)
2. Rossion, B., Gauthier, I.: How Does the Brain Process Upright and Inverted Faces? Behav. Cogn. Neurosci. Rev. **1**, 63–75 (2002)
3. Bukach, C.M., Gauthier, I., Tarr, M.J.: Beyond faces and modularity: The power of an expertise framework. Trends Cogn. Sci. **10**, 159–166 (2006)
4. Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., Boyes-Braem, P.: Basic objects in natural categories. Cogn. Psychol. **8**, 382–439 (1976)
5. Tanaka, J.W., Taylor, M.: Object categories and expertise: Is the basic level in the eye of the beholder? Cogn. Psychol. **23**, 457–482 (1991)
6. Chase, W.G., Simon, H.A.: Perception in chess. Cogn. Psychol. **4**, 55–81 (1973)
7. Charness, N.: Components of skill in bridge. Canad J Psychol. **33**, 1–16 (1979)
8. Beal, A.L.: The skill of recognizing musical structures. Memory Cogn. **13**, 405–412 (1985)
9. Egan, D.E., Schwartz, B.J.: Chunking in recall of symbolic drawings. Mem. Cogn. **7**, 149–158 (1979)
10. Nosofsky, R.M.: Attention and learning processes in the identification and categorization of integral stimuli. J Exp Psychol Learn, Mem. Cogn. **13**, 87–108 (1987)
11. Zhang, L., Cottrell, G.W.: A computational model which learns to selectively attend in category learning. Proceedings of the 2005 International Conference on Development and Learning, vol. 19. pp. 195–200
12. Johansen, M.K., Palmeri, T.J.: Are there representational shifts during category learning? Cogn. Psychol. **45**, 482–553 (2002)
13. Sowden, P.T., Davies, I.R.L., Rolings, P.: Perceptual learning of the detection of features in X-ray images: A functional role for improvements in adults' visual sensitivity? J. Exp. Psychol.: Human Percep. Perform. **26**, 379–390 (2000)
14. Dror, I.E., Peron, A.E., Hind, S.L., Charlton, D.: When emotions get the better f us: The effect of contextual top-down processing on matching fingerprints. Appl. Cogn. Psychol. **19**, 799–809 (2005)
15. Dror, I.E., Charlton, D.: Why experts make errors. J. Forensic Identif. **56**, 600–616 (2006)
16. Dror, I.E., Charlton, D., Peron, A.E.: Contextual information renders experts vulnerable to making erroneous identifications. Forensic Sci. Int. **156**, 74–78 (2006)
17. Dror, I., Rosenthal, R.: Meta-analytically Quantifying the Reliability and Biasability of Forensic Experts. J. Forensic Sci. **53**, 900–903 (2008)

Latex Finger

- ▶ Fingerprint Fake Detection

Law Enforcement

Ken Moses
Forensic Indentification Services, San Franscisco, USA

Synonyms

Criminal law enforcement; Law enforcement agency; Police law enforcement

Definition

Law enforcement systems record information about individuals, both biometric data and demographic data, that is used for quickly identifying criminals and their acts.

Introduction

The identification of criminals and repeat offenders has been one of the primary missions of police agencies since their creation in the nineteenth century. The industrial revolution led to mass migration to cities where anonymity gave safe harbor to lawbreakers. Newspapers fanned both real and perceived threats to public safety and placed persistant demands upon the police to solve crimes. Laws were enacted that dealt the

harshest penalties to repeat offenders encouraging lawbreakers to adopt a different name (known as an alias) with each new arrest.

Answering the call for effective systems of identification, law enforcement agencies initiated various systems to maintain records at the end of the nineteenth century. One of the earliest techniques utilized the new science of photography to take pictures or mug shots at the time of arrest. These mugs shots were organized into albums, called Rogues Galleries. The method of recording these photographs was adopted and standardized worldwide to include a front and profile view of the subject. General physical descriptions, including scars, marks, and tattoos accompanied the photographs. Today, Rogues Galleries have evolved into large digitized databases of mug shots from which investigators can select possible suspects based on victims' descriptions, or compile "six packs" of persons sharing similar descriptions for use in photo line-ups [1]. Automated facial recognition would be a natural extension of the mug shot systems, and research is underway to build that capability into the system that would be most valuable in video surveillance cases.

In 1880, Alphonse Bertillon, a clerk in the Paris police, devised an anthropomorphic system of identification based upon 18 body measurements utilizing specialized rulers and calipers for precision. From these measurements, Bertillon derived a numeric value or "portrait parle," that was supposedly unique to each person. Bertillon's Portrait Parle met with great success in identifying repeat offenders regardless of what names they had assumed at the time of arrest. Within a decade, this manual biometric system was in use in police agencies worldwide (Fig. 1).

Law Enforcement. **Figure 1** Early police biometrics circa. 1906 Bertillon "Portrait Parle" record.

During this period, a new type of biometric was being explored by William Herschel in India and Dr. Henry Faulds in Japan. In 1892, Francis Galton, the nephew of Charles Darwin, published *Finger Prints* [2], a book that became the standard reference on the subject. Sir Edward Henry, police commissioner in Bengal India, devised a classification system that allowed for easy search and retrieval of fingerprints. The Henry System was later adapted by Scotland Yard in 1901. For several years, police agencies around the world utilized both the Bertillon Portrait Parle and fingerprinting systems side by side until fingerprints won out and Portrait Parle was largely abandoned. Fingerprints proved easier to record, file, and retrieve, and they could be used to solve crimes when they were inadvertently left by criminals at the scene of crime [3].

Today, as with mug shots, fingerprints are usually digitally recorded, stored, and automatically searched by Automated Fingerprint Identification Systems (AFIS) [4]. AFIS has become the backbone of criminal identification and is the most widely used biometric system in the world. Standards have been adopted for the digital storage and transmission of fingerprint images [5].

In *Finger Prints*, Galton highlighted a number of characteristics of each fingerprint that are used to uniquely identify it. Modern automated systems use some of these Galton characteristics in extracting information from the print (Fig. 2).

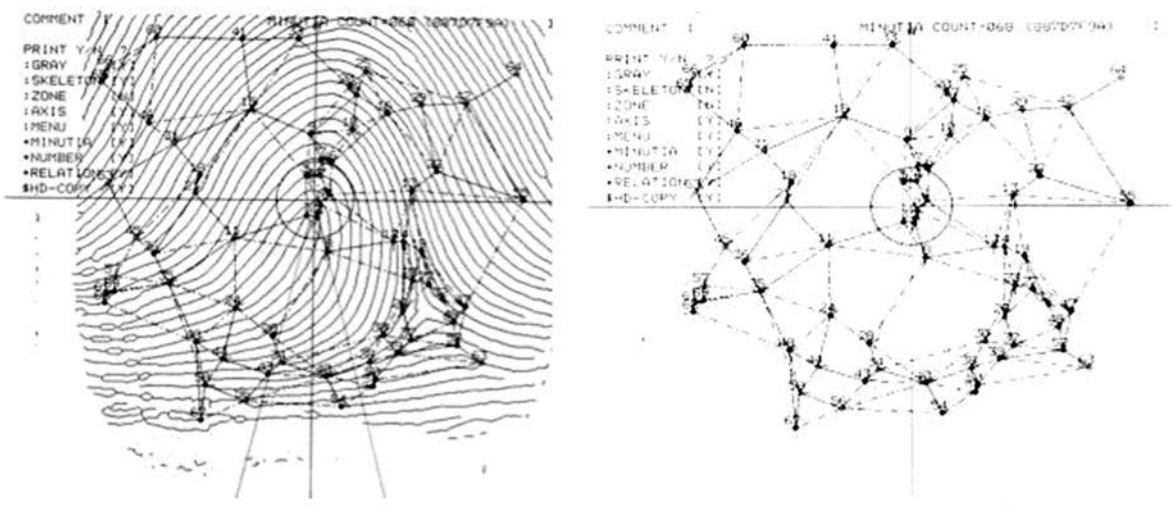
Fingerprint Functions in Law Enforcement

Law enforcement AFIS systems are composed of two interdependent subsystems: the tenprint or criminal ID subsystem, and the latent or criminal investigation subsystem. Each subsystem operates with a considerable amount of autonomy, and both are vital to public safety.

The Tenprint subsystem is tasked with identifying sets of inked or scanned fingerprints incident to an arrest or citation or as part of the applicant process to determine whether or not the person has an existing record. An automated tenprint search usually involves a search of only two or four fingers. Submitted fingerprints by and large have sufficient clarity and detail to make searching of more than two fingers redundant. At present the AFIS can return a search of one million records in under a minute.

The latent print or criminal investigation subsystem is tasked with solving crimes through the identification of latent prints developed from the scene of crime. Terminals used within the latent subsystem are often specialized to accommodate the capture and digital enhancement of individual latent prints (Fig. 3).

The search of a latent print is more tedious and time consuming than a tenprint search. Latent prints are often fragmentary and of poor image quality. Minutiae features must be reviewed by a latent print



Law Enforcement. **Figure 2** Computer plotting of minutiae or Galton details from a fingerprint.

Law Enforcement. **Figure 3** Latent fingerprint workstation.

examiner one by one before the search even begins. Once the search is launched, the AFIS searches the database and returns a respondent list of the closest matches [6–8]. The fingerprint examiner then compares each candidate against the search print.

The identification of latent fingerprints from the scenes of crime has been a tremendous boon to law enforcement agencies. AFIS identifications, or hits, are responsible for solving hundreds of thousands of crimes each year throughout the world. In 2005, in the United States alone, 50,000 cases were cleared by AFIS hits.

References

1. The Criminal Investigative Process; Rand, R-1776-DOJ; October (1975)
2. Galton, F.: Finger Prints. Macmillan, London (1892)
3. Cole, S.: Suspect Identities. Harvard University Press, Cambridge, MA (2001)
4. Lee, H.C., Gaensslen, R.E.: Advances in Fingerprint Technology, 2nd edn. CRC Press, West Palm Beach, FL (2001)
5. Data Format for the Interchange of Fingerprint, Facial, & Scar Mark & Tattoo (SMT) Information (ANSI/NIST-ITL 1-2000)
6. Komarinski, P.: Automated Fingerprint Identification Systems. Elsevier, Amsterdam (2005)
7. Ratha, N., Bolle, R.: Automated Fingerprint Recognition Systems. Springer, New York (2004)
8. Maltoni, D., et al.: Handbook of Fingerprint Recognition. Springer, New York (2003)

Law Enforcement Agency

► Law Enforcement

LBP (Local Binary Pattern)

► Local Image Filters

LCN DNA/Low Template Level

Low template level or as formally referred to as Low Copy Number (LCN) DNA refers to the analysis of trace amounts of DNA (typically less than 100 pg). As LCN is an ultrasensitive technique, there are artefacts that one must be aware (allele drop-in and drop-out). It requires careful interpretation and very strict protocols to avoid contamination.

► Forensic DNA Evidence

LDA (Linear Discriminant Analysis)

Linear discriminant analysis is defined as an orthogonal linear transformation that best separate two or more classes of objects. It finds the set of the projection vectors which maximize the ratio of between-class scatter against within-class scatter. The resulting combination may be used for classification.

► Face Variation
► Illumination Compensation
► Linear Dimension Reduction

LED (Light Emission Diode)

LED refers to an electronic component, Light Emission Diode, which emits light when traversed by a current of certain entity.

A light-emitting diode, also called an LED, is a semiconductor diode that converts electric energy into electromagnetic radiation at a visible and near infrared frequencies when its p–n junction is forward biased.

► Biometric Sensor and Device, Overview
► Face Recognition, Near-infrared

Lighting Compensation

► Illumination Compensation

Lighting Model

The realism of a computer generated visual scene depends on the extent to which lighting phenomena are mimicked by the lighting model used. In Computer Graphics, the lighting model refers to the choice of reflectance function used. However, in Computer Vision, it may also include the type of light source(s) and of normals used, as these have a considerable effect on the synthesis and on the analysis by synthesis results. A reflectance function is a mathematical function that describes how light is reflected from an object.

► Face Sample Synthesis

Likelihood Ratio Test

The likelihood ratio test results from simplifying the ratio between the probability that the sensor's observation resulted from the positive hypothesis (genuine user) divided by the probability that the sensor's observation resulted from the negative hypothesis (imposter). The objective is to improve the ratio by making the probability distributions separate as much as possible. If the distributions are far apart, the probability that the observation came from the genuine user increases, and the probability that the observation came from the imposter decreases. This increases the ratio improving the test.

► Fusion, Decision Level

Limbus

The limbus is the outer boundary of the annular iris structure, where the iris meets the white portion of the eye, the sclera.

► Iris Image Data Interchange Formats, Standardization

Linear Dimension Reduction

Wei-Shi Zheng[1], J. H. Lai[1], Pong C. Yuen[2]
[1]Sun Yat-sen University, Guangzhou, P.R. China
[2]Hong Kong Baptist University, Hong Kong, China

Synonym

Linear feature extraction

Definition

Linear dimension reduction technique reduces the dimension of biometric data using a linear transform. The linear transform is always learned by optimization of a criterion. Biometric data are then projected on to the range space of this transform. Subsequent processing is then performed in that lower-dimensional space.

Introduction

In biometrics, data are invariably represented in vectors and the dimensionality is consistently very high. It would be computationally expensive to process them directly by using many algorithms. Moreover, it is sometimes desirable to exact robust, informative or discriminative facts contained in the data. For these reasons, a lower dimensional subspace is found such that the most important part of the data is retained for linear representation. Among the techniques for learning such subspace, linear dimension reduction methods are popular.

Given a set of N data samples $\{\mathbf{x}_1,\ldots,\mathbf{x}_N\}$, where $\mathbf{x}_i \in \Re^n$. Linear dimension reduction technique finds a linear transform matrix $\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_\ell)$ such that data are projected on to the range space $span\{\mathbf{w}_1,\ldots,\mathbf{w}_\ell\}$ by

$$\mathbf{y}_i = \mathbf{W}^T\mathbf{x}_i, \tag{1}$$

where ^{T} denotes the transpose and $\mathbf{y}_i$ is the representation of $\mathbf{x}_i$ in the lower-dimensional space.

The linear dimension reduction technique is equal to extraction of linear features $\mathbf{w}_1,\ldots,\mathbf{w}_\ell$. Different linear dimension reduction techniques lie in different goals of the information retained by these linear features. Generally, they can be categorized into three classes, which address the linear dimension problem in the ▶ unsupervised, ▶ supervised, and ▶ semi-supervised cases respectively. Some representative algorithms of these classes are described here.

When biometric data are represented in a matrix form, linear dimension reduction techniques can also be extended. This kind of extension is called the two-dimensional linear dimension reduction technique.

Unsupervised Linear Dimension Reduction

Unsupervised linear dimension reduction aims at the extraction of reconstructive features. Biometric data are then linearly approximated in a lower-dimensional subspace spanned by the reconstructive features. Among the popular linear algorithms used in biometrics for dimension reduction, the Principal Component Analysis (PCA) [1] is the most representative one. The PCA finds a lower-dimensional space that preserves the greatest variations of data, that is, an optimal transform is learned by maximization of the following criterion:

$$\mathbf{W}_{opt} = \arg\max_{\mathbf{W}^T\mathbf{W}=\mathbf{I}} \; trace(\mathbf{W}^T\mathbf{C}_t\mathbf{W}), \tag{2}$$

where $\mathbf{C}_t$ is the total-class covariance matrix defined by

$$\mathbf{C}_t = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i-\mathbf{u})(\mathbf{x}_i-\mathbf{u})^T,\ \mathbf{u} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i. \tag{3}$$

Eigen-decomposition is used to find $\mathbf{w}_1,\ldots,\mathbf{w}_\ell$, which are eigenvectors corresponding to the largest ℓ eigenvalues. However, preserving the largest eigenvectors may not be the best strategy for other applications of the PCA, such as recognition and clustering. In these cases, the selection of proper principal components may be useful.

Features extracted by the PCA are statistically uncorrelated, but they are not ensured to be statistically independent. If it is assumed that data are approximately linear representations of some independent sources, then it is useful to find these independent components for representation of data in an intrinsic subspace. The Independent Component Analysis (ICA) [2] pursues features in this aspect. Let $\mathbf{X} = [\mathbf{x}_1,\ldots,\mathbf{x}_N]$ be the data matrix. In the ICA, data matrix $\mathbf{X}$ is approximated by a multiplication of mixing matrix $\mathbf{A}$ and matrix $\mathbf{S}$ which consists of independent components as follows

$$\mathbf{X} \approx \mathbf{AS}. \tag{4}$$

Several approaches have been proposed for the estimation of the ICA and the most popular of these is FastICA [2].

In biometric learning, especially facial image analysis, has been experimentally demonstrated that the extraction of localized features is useful for recognition and for the interpretation and understanding of the structure of data as well. Unlike the PCA and the ICA, non-negative matrix factorization (NMF) [3] is a novel technique for this purpose. In the NMF, data are approximated only by additive combination of non-negative components. The NMF, therefore, finds two non-negative matrices, namely, the component matrix $\mathbf{W}_{opt}$ and the coefficient matrix $\mathbf{H}_{opt}$, which are minimums of the following criterion

$$(\mathbf{W}_{opt}, \mathbf{H}_{opt}) = \arg\min_{(\mathbf{W},\mathbf{H})} ||\mathbf{X} - \mathbf{WH}||_F^2, \quad s.t.\ \ \mathbf{W} \geq 0\ \&\ \mathbf{H} \geq 0. \tag{5}$$

where $||\cdot||_F$ is the Frobenius norm. Other than using the Euclidean distance, the following criterion based on the generalized Kullback-Leibler divergence that is lower bounded by Zoro is also popular

$$(\mathbf{W}_{opt}, \mathbf{H}_{opt}) = \arg\min_{(\mathbf{W},\mathbf{H})} \sum_{i,j} \mathbf{X}_{ij} \log \frac{\mathbf{X}_{ij}}{(\mathbf{WH})_{ij}} - \mathbf{X}_{ij} + (\mathbf{WH})_{ij}, \quad s.t.\ \ \mathbf{W} \geq 0\ \&\ \mathbf{H} \geq 0. \tag{6}$$

Closed forms of $\mathbf{W}_{opt}$ and $\mathbf{H}_{opt}$ in the NMF are difficult to obtain. At present, the multiplicative update method [4] has been widely used for finding a locally optimal solution.

The former three methods do not take the geometric relationship between data into account. Assume data are actually distributed on a manifold, it is natural that the neighboring data in the input space would also be close to each other when they are represented in the lower rank subspace. The locality preserving projection (LPP) [5] is a typical algorithm developed for this purpose. It learns an optimal transform $\mathbf{W}_{opt}$ such that the locality relationship between data in the input data space are preserved after dimension reduction. The cost function of a transform $\mathbf{W}$ for locality preserving is modeled by

$$\frac{1}{2}\sum_{i,j} g_{ij}||\mathbf{y}_i - \mathbf{y}_j||^2 = \frac{1}{2} trace(\mathbf{W}^T \sum_{i,j} g_{ij} \cdot (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}),$$

where g_{ij} the affinity between samples $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined by

$$g_{ij} = \begin{cases} 1 & \text{if } \mathbf{x}_i \in \mathcal{N}_s(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in \mathcal{N}_s(\mathbf{x}_i), \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

and $\mathcal{N}_s(\mathbf{x}_i)$ is the set of s nearest neighbors of data $\mathbf{x}_i$. Then the optimal transform explored by the LPP is learned as follows

$$\mathbf{W}_{opt} = \arg\min_{\mathbf{w}_i^T \mathbf{XDX}^T \mathbf{w}_i = 1} trace(\mathbf{W}^T \mathbf{S}_p \mathbf{W}), \tag{8}$$

where $\mathbf{D} = diag(\sum_j g_{1j}, \ldots, \sum_j g_{Nj})$, $\mathbf{S}_p = \mathbf{XLX}^T$ and $\mathbf{L}$ is termed the Laplacian matrix (He and Niyogi [5]) formulated by

$$\mathbf{L} = \mathbf{D} - \mathbf{G}, \quad \mathbf{G}_{ij} = g_{ij}. \tag{9}$$

Supervised Linear Dimension Reduction

PCA, the ICA, the NMF and the LPP are methods for learning unlabeled data. Since the features extracted by these methods are not discriminative, they may not be useful for recognition. When class labels of data are known, supervised techniques for dimension reduction can be developed by additionally using label information. Suppose data $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are drawn from L classes, namely, classes $C_1, \ldots, C_L$, and N_k be the number of samples of class C_k, the supervised dimension reduction technique then finds a transform such that data of the same class are close to each other while data of different classes are scattered as far as possible. A popular and widely applied technique is (Generalized) Fisher's linear discriminant analysis (LDA) [6]. The LDA finds a lower-dimensional subspace in which the ratio between between-class variance and within-class variance is maximized. That is, a discriminative linear transform $\mathbf{W}_{opt}$ is the maximum of the following criterion

$$\mathbf{W}_{opt} = \arg\max_{\mathbf{W}} \frac{trace(\mathbf{W}^T \mathbf{S}_b \mathbf{W})}{trace(\mathbf{W}^T \mathbf{S}_w \mathbf{W})}, \tag{10}$$

where $\mathbf{S}_b$ and $\mathbf{S}_w$ are between-class and within-class covariance matrices respectively, which are defined as follows:

$$\mathbf{S}_w = \frac{1}{N}\sum_{k=1}^{L} \sum_{\mathbf{x}_i \in C_k} (\mathbf{x}_i - \mathbf{u}_k)(\mathbf{x}_i - \mathbf{u}_k)^T, \quad \mathbf{u}_k = \frac{1}{N_k}\sum_{\mathbf{x}_i \in C_k} \mathbf{x}_i,$$

$$\mathbf{S}_b = \frac{1}{N}\sum_{k=1}^{L} N_k (\mathbf{u}_k - \mathbf{u})(\mathbf{u}_k - \mathbf{u})^T.$$

To obtain the optimum of a criterion [10], the eigenvectors of $\mathbf{S}_w{}^{-1}\mathbf{S}_b$ have to be determined. As dimensionality of data is high and training samples are in invariably limited in biometric learning, within-class covariance matrix $\mathbf{S}_w$ will be singular. This kind of singularity problem is referred to as the small sample size problem in the LDA. Techniques used to address this problem include, PCA+LDA [7], the Null-space LDA (N-LDA) (Chen et al. [8]), and the regularized LDA [9]. In the PCA+LDA, LDA is performed in a principal component subspace in which within-class covariance matrix is of full rank. In the N-LDA, the nullspace of within-class covariance matrix $\mathbf{S}_w$ is first learned; then data are projected on to that subspace; finally the discriminant transform is found for

maximization of the variance of between-class data. Though both the PCA+LDA and the N-LDA are lower-rank methods that first project data onto a lower-dimensional subspace before implementation of any discriminant processing, they are different in that $\mathbf{S}_w$ in the PCA+LDA will be of full rank after dimension reduction by the PCA while $trace(\mathbf{W}^T\mathbf{S}_w\mathbf{W}) = 0$ after dimension reduction in the N-LDA. Unlike the lower-rank approach, in the regularized LDA (R-LDA), a regularized term, such as $\alpha\mathbf{I}$ where $\alpha > 0$, is added to matrix $\mathbf{S}_w$ in Eq. (10), so that Fisher criterion is well-conditioned. Regularization is a straightforward way to solve this singularity problem, but the regularized parameter will have a significant impact on the performance of the R-LDA.

An alternative way to maximize between-class variance as well as minimize within-class variance can be modeled as follows:

$$\mathbf{W}_{opt} = \arg\max_{\mathbf{W}} \; trace(\mathbf{W}^T\mathbf{S}_b\mathbf{W}) - \lambda \cdot trace(\mathbf{W}^T\mathbf{S}_w\mathbf{W}), \tag{11}$$

where $\lambda > 0$. The advantage of using this model is that computation of the inverse of $\mathbf{S}_w$ is not required, but how to determine the importance weight λ could be a problem. This criterion is known as the maximum margin criterion [10] when $\lambda = 1$.

At times, unsupervised linear dimension reduction techniques can be used as a preprocessing step before applying supervised techniques. Methods driven in this way are two-step dimension reduction techniques. The PCA+LDA is typically in line with this approach. In addition, though supervised methods are always preferred for recognition, it is difficult to ascertain which kind of linear dimension reduction technique is the best. For example, the PCA may be better than the LDA for face recognition if the number of training samples for each class is small [11].

Semi-Supervised Linear Dimension Reduction

Linear dimension reduction for partially labeled data would be highly demanded for large scale problems, since labeling data is an expensive task. Therefore, it is desirable to utilize unlabeled data for extraction of supervised features for dimension reduction. Among the developed techniques to achieve this goal, a special regularized LDA for performing linear dimension reduction on partially labeled data can be formulated as follows [12]:

$$\mathbf{W}_{opt} = \arg\max_{\mathbf{W}} \frac{trace(\mathbf{W}^T\mathbf{S}_b\mathbf{W})}{trace(\mathbf{W}^T(\mathbf{S}_t + \beta\mathbf{S}_p)\mathbf{W})}, \tag{12}$$

where $\mathbf{S}_t = \mathbf{S}_w + \mathbf{S}_b$ and $\beta > 0$. In this criterion, labeled data are used to estimate supervised class covariance information, and the effect of unlabeled data is reflected by the Laplacian term $\mathbf{S}_p$. The underusing idea is that labeled data are separated in the same way as done in the LDA while the neighboring data including unlabeled data are nearby after dimension reduction.

Two-Dimensional Linear Dimension Reduction Techniques

Many well-known linear dimension reduction techniques assume that input patterns are represented in vectors. However, biometric data are captured in images, and the dimensionality is very high when they are reshaped into vectors. Unlike traditional techniques, some linear dimension reduction techniques are developed by performing linear transformation directly on matrix form data, such as image matrices. This is advantageous in tackling large scale problems. Suppose $\mathbf{X}_1, \ldots, \mathbf{X}_N$ are the corresponding matrix representations of vector form data $\mathbf{x}_1, \ldots, \mathbf{x}_N$, then a transform $\mathbf{W}$ for dimension reduction of $\mathbf{X}_i$ would perform as follows

$$\mathbf{Y}_i = \mathbf{W}^T\mathbf{X}_i, \tag{13}$$

where $\mathbf{Y}_i$ is the representation after dimension reduction. Among the developed techniques are two representative algorithms: two-dimensional principal component analysis (2D-PCA) [13] and two-dimensional linear discriminant analysis (2D-LDA) [14]. The vector-based PCA and LDA are also referred to as the 1D-PCA and 1D-LDA respectively.

The main difference between 2D-PCA and 1D-PCA as well as between 2D-LDA and 1D-LDA lies in their different means of covariance matrix estimation. In two-dimensional linear reduction techniques, the covariance matrices are calculated directly based on data represented in matrix form. Apart from this, the main

ideas of 2D-PCA and 2D-LDA are almost similar to those of 1D-PCA and 1D-LDA respectively. More specifically, 2D-PCA learns the optimal transform via the following criterion

$$\mathbf{W}_{opt}^{2d} = \arg\max_{\mathbf{W}^T\mathbf{W}=\mathbf{I}} trace(\mathbf{W}^T\mathbf{C}_t^{2d}\mathbf{W}), \qquad (14)$$

where

$$\mathbf{C}_t^{2d} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{X}_i - \mathbf{U})(\mathbf{X}_i - \mathbf{U})^T, \quad \mathbf{U} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_i.$$

For 2D-LDA, the criterion is modified in a similar manner as follows:

$$\mathbf{W}_{opt}^{2d} = \arg\max_{\mathbf{W}} \frac{trace(\mathbf{W}^T\mathbf{S}_b^{2d}\mathbf{W})}{trace(\mathbf{W}^T\mathbf{S}_w^{2d}\mathbf{W})}, \qquad (15)$$

where

$$\mathbf{S}_w^{2d} = \frac{1}{N}\sum_{k=1}^{L}\sum_{\mathbf{X}_i\in C_k}(\mathbf{X}_i - \mathbf{U}_k)(\mathbf{X}_i - \mathbf{U}_k)^T,$$

$$\mathbf{U}_k = \frac{1}{N_k}\sum_{\mathbf{X}_i\in C_k}\mathbf{X}_i.$$

$$\mathbf{S}_b^{2d} = \sum_{k=1}^{L}\frac{N_k}{N}(\mathbf{U}_k - \mathbf{U})(\mathbf{U}_k - \mathbf{U})^T.$$

The above 2D-PCA and 2D-LDA are unilateral, which means only one transform matrix multiplied on one side of the data matrix is available. To overcome this limitation, there are generalizations such as bilateral 2D-PCA [15] and bilateral 2D-LDA [16], which learn transform matrices $\mathbf{W}_l$ and $\mathbf{W}_r$ on both sides of the data matrix and perform dimension reduction for data $\mathbf{X}_i$ as follows

$$\mathbf{Y}_i = \mathbf{W}_l^T\mathbf{X}_i\mathbf{W}_r. \qquad (16)$$

However, it would be difficult to obtain a closed solution for bilateral techniques, and alternating optimization methods are used for finding a locally optimized solution.

In general, two-dimensional linear dimension reduction techniques gain lower cost of computation, and 2D-LDA in particular avoids the small sample size problem. However, it is cannot be definitely said that two-dimensional techniques are better. An insightful analysis and extensive comparisons between 2D-LDA and 1D-LDA have been made by [17]. Besides the theoretical comparisons, Zheng et al. noted that 2D-LDA is not always better than 1D-LDA when the number of training samples for each class or the number of extracted features is small.

Summary

Linear dimension reduction is an important step for processing of biometric data. It is equal to extraction of a set of linear feature vectors, which span a lower-dimensional subspace. Linear techniques for finding the most robust, informative and discriminative information are fundamental technologies in pattern recognition. Besides, the development of new linear techniques for large scale problems in biometric learning is an important topic.

Related Entries

- ▶ Feature Extraction
- ▶ Kernel Methods
- ▶ LDA (Linear Discriminant Analysis)
- ▶ Manifold Learning
- ▶ PCA (Principal Component Analysis)

References

1. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
2. Hyvärinen, A., Oja, E.: Independent component analysis: algorithms and applications. Neural Networks **13**, 411–430 (2000)
3. Lee, D.D., Seung, H.S.: Learning the parts of objects by non-negative matrix factorization. Nature **401**, 788–791 (1999)
4. Lee, D.D., Seung, H.S.: Algorithms for non-negative matrix factorization. In: Advances in Neural Information Processing Systems, Denver, Colorado pp. 556–562 (2000)
5. He, X., Niyogi, P.: Locality preserving projections. In: Advances in Neural Information Processing Systems, Vancauver, Canada pp. 153–160 (2003)
6. A., F.R.: The Use of Multiple Measures in Taxonomic Problems. Ann. Fisher, RA Eugenics **7,** 179–188 (1936)
7. Belhumeour, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Analy. Mach. Intell. **19**(7), 711–720 (1997)

8. Chen, L.F., Liao, H.Y.M., Ko, M.T., Lin, J.C., Yu, G.J.: A new LDA-based face recognition system which can solve the small sample size problem. Pattern Recogn. **33**, 1713–1726 (2000)
9. Webb, A.R. (ed.): Statistical Pattern Recognition, 2nd ed. 2002. Wiley, England (2002)
10. Li, H., Jiang, T., Zhang, K. Efficient and robust feature extraction by maximum margin criterion. IEEE Trans. Neural Networks **17**(1), 157–165 (2006)
11. Martínez, A.M., Kak, A.C.: PCA versus LDA. IEEE Trans. Pattern Analy. Mach. Intell. **23**(2), 228–233 (2001)
12. Cai, D., He, X., Han, J.: Semi-supervised Discriminant Analysis. In: IEEE Int. Conf. Comput. Vis. Rio de Janeiro, Brazil (2007)
13. Yang, J., Zhang, D., Frangi, A.F., Yang, J.y.: Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans. Pattern Analy. Mach. Intell. **26**(1), 131–137 (2004)
14. Xiong, H., Swamy, M.N.S., Ahmad, M.O.: Two-dimensional FLD for face recognition. Pattern Recogn. **38**, 1121–1124 (2005)
15. Kong, H., Li, X., Wang, L., Teoh, E.K., J.-G., W., Venkateswarlu, R.: Generalized 2D principal component analysis. In: IEEE International Joint Conference on Neural Networks, vol. 1, pp. 108–113. Oxford, UK (2005)
16. Ye, J.P., Janardan, R., Li, Q.: Two-dimensional linear discriminant analysis. In: Advances in Neural Information Processing Systems, pp. 1569–1576 (2004)
17. Zheng, W.S., Lai, J.H., Li, S.Z.: 1D-LDA versus 2D-LDA: when is vector-based linear discriminant analysis better than matrix-based? Pattern Recogn. **41**(7), 2156–2172 (2008)

Linear Feature Extraction

► Linear Dimension Reduction

Linearly Symmetric Image

An image described by $f(x, y)$ is linearly symmetric if its isocurves have a common direction, i.e., there exists a 1D function $h(\tau)$ such that $f(x, y) = h(kxx + kyy)$ for a certain (constant) direction vector $k = (kx, ky)T$.

► Fingerprint Features

Lip Movement Recognition

Petar S. Aleksic
Google Inc., New York, NY, USA

Synonyms

Audio–visual-dynamic speaker recognition; Visual-dynamic speaker recognition

Definition

Lip movement recognition is a speaker recognition technique, where the identity of a speaker is determined/verified by exploiting information contained in dynamics of changes of visual features extracted from the mouth region. The visual features usually consist of appropriate representations of the mouth appearance and/or shape. This dynamic visual information can also be used in addition to the acoustic information in order to improve the performance of audio-only speaker recognition systems and increase their resilience to spoofing, therefore giving rise to audio–visual-dynamic speaker recognition systems.

Introduction

Speech contains information about identity, emotion, location, as well as linguistic information, and plays a significant role in the development of human computer interaction (HCI) systems, including speaker recognition systems. However, audio-only systems can perform poorly even at typical acoustic background signal-to-noise ratio levels (−10 to 20 dB). In addition, they can be sensitive to microphone types (desktop, telephone, etc.), acoustic environment (car, plane, factory, babble, etc.), channel noise (telephone lines, VoIP, etc.), or complexity of the scenario (speech under stress, Lombard speech, whispered speech).

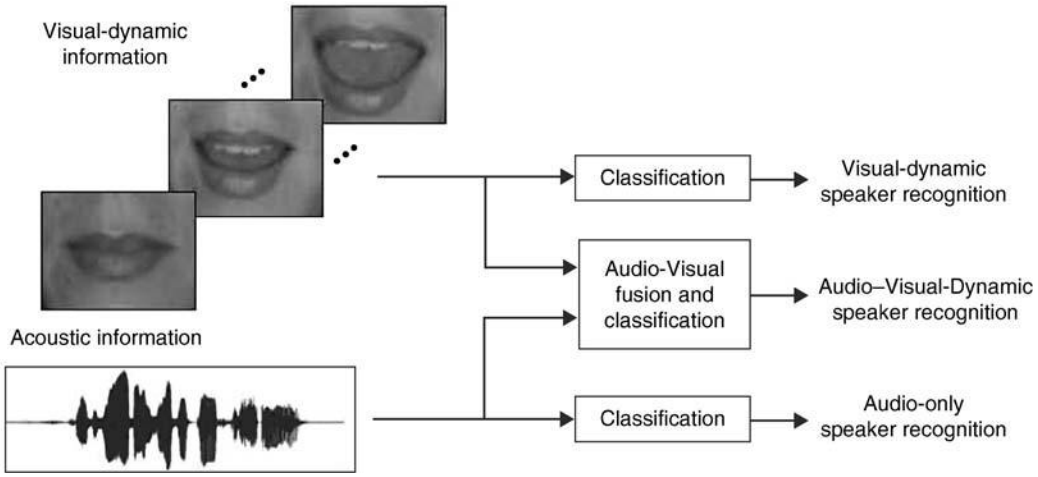
Similarly, the visual modality can be exploited to improve HCI [1, 2]. Visual facial features extracted from the mouth region are both correlated to the produced audio signal [3] and also contain complementary information to it [4]. Lip movement recognition systems, also known as visual-dynamic speaker

recognition systems, exploit information contained in dynamics of changes of visual features extracted from video sequences of the mouth area (see Fig. 1). Although speaker recognition systems that utilize only visual-dynamic information can achieve high recognition rates, visual-dynamic information is typically utilized together with acoustic information to improve speaker recognition performance, resulting in audio–visual-dynamic speaker recognition systems (see Fig. 1). There is a number of audio–visual fusion techniques [5], that combine audio and visual information to achieve higher recognition performance than both audio-only and visual-only systems. The use of visual information improves speaker recognition performance even in noise-free environments [5, 6]. The potential for such improvements is even greater in acoustically noisy environments, since visual features are typically much less affected by acoustic noise than the acoustic features. In addition, audio-only, as well as static-image-based (face recognition) person recognition systems are vulnerable to impostor attacks (spoofing), if the impostor possesses a photograph and/or speech recordings of the client. However, it is significantly more difficult for an impostor to impersonate both acoustic and dynamical-visual (lip movements) information simultaneously. The main advantage of audio–visual-dynamic speaker recognition systems lies in their robustness, since each modality can provide independent and complementary information and therefore prevent performance degradation due to the noise present in one or both of the modalities, as well as increase resilience to spoofing. Another advantage of audio–visual speaker recognition systems is that they use unobtrusive and low-cost methods for extracting biometric features, thus enabling natural speaker recognition and reducing inconvenience.

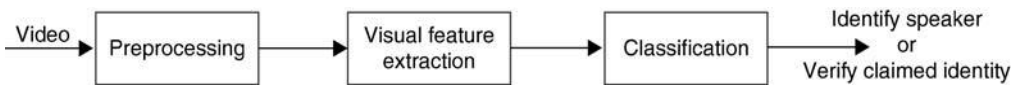
Operation of a Lip-Movement Recognition System

The block diagram of a visual-dynamic speaker recognition system is shown in Fig. 2. It consists of three main blocks, corresponding to *preprocessing*, *visual feature extraction*, and *classification* phases.

In the preprocessing phase, detection and tracking of the face and the important facial features in the input video are performed. These tasks represent challenging problems, especially in cases where the background, head pose, and lighting are varying [7]. The most commonly used face detection techniques use statistical modeling of the face appearance to obtain a classification



Lip Movement Recognition. **Figure 1** Speaker recognition systems.



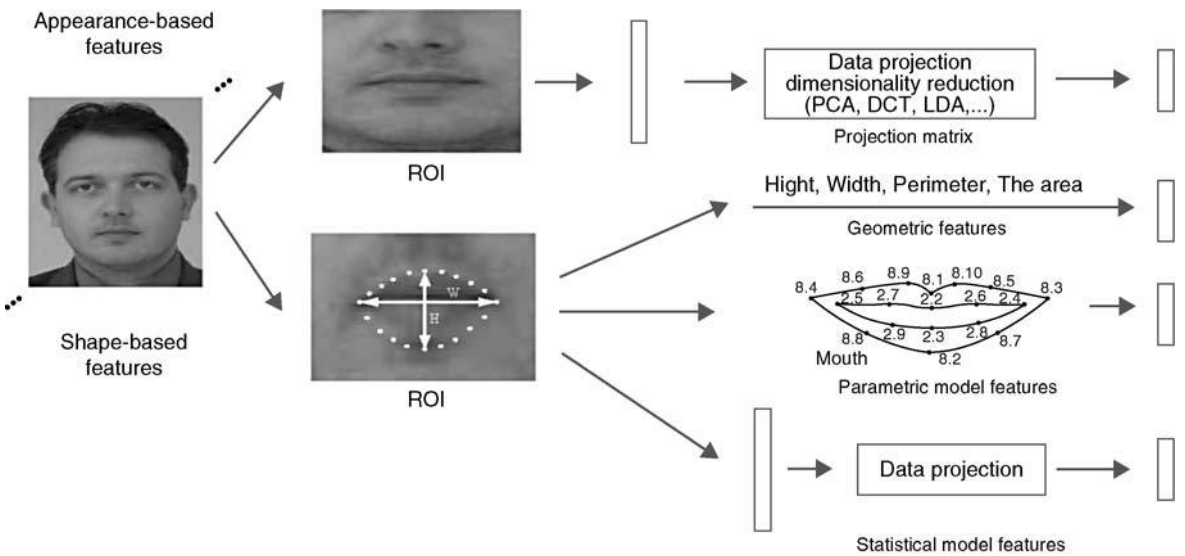
Lip Movement Recognition. **Figure 2** Block diagram of a visual-dynamic speaker recognition system.

of image regions into face and non face classes. These regions are typically represented as vectors of grayscale or color image pixel intensities. They are often projected onto lower dimensional spaces, and are defined over a number of possible locations, scales, and orientations in the video frame. One or more techniques, such as artificial neural networks (ANNs), support vector machines (SVMs), Gaussian mixture models (GMMs), clustering algorithms, and linear discriminants, are usually utilized to classify these regions. Typically, face detection goes hand-in-hand with tracking in which the temporal correlation is taken into account.

After successful face detection, facial features, such as mouth corners, eyes, nostrils, and chin, are detected by utilizing prior knowledge of their relative position on the face to simplify the search. If color information is available, it can be utilized to detect certain facial features (especially lips). Subsequently, the mouth region-of-interest (ROI), containing useful visual information is extracted. The ROI is typically a rectangle containing the mouth, possibly including larger parts of the lower face, such as the jaw and cheeks (see Fig. 3). The normalization is usually performed with respect to head-pose information and lighting. The ROI can also be extended into a three-dimensional rectangle, containing adjacent frame ROIs, thus capturing dynamic visual information.

In the visual feature extraction phase, relevant visual features from the ROI are extracted. Investigating choice of visual features for speaker recognition is a relatively new research topic. The various sets of the choice facial features proposed in the literature are generally grouped into three categories [8]: (1) *appearance-based features*; (2) *shape-based features*; and (3) features that are a combination of both appearance and shape features [5]. The appearance-based visual features are extracted from the ROI using image transforms, such as principal component analysis (PCA,) generating "eigenlips" [9], the discrete cosine transform (DCT) [10], the discrete wavelet transform (DWT) [10], linear discriminant analysis (LDA) [2, 5, 11], Fisher linear discriminant (FLD), etc. (see Fig. 3). These transforms are applied on a feature vector, created by ordering the grayscale pixel values inside the ROI, and are commonly applied in series to cope with the "curse of dimensionality" problem.

Shape-based visual mouth features are divided into *geometric*, *parametric*, and *statistical* (see Fig. 3). With shape-based features it is assumed that most of the information is contained in the shape of the speaker's lips [5, 12]. Hence, such features achieve a compact representation of mouth images using low-dimensional vectors. Geometric features, such as the height, width, perimeter of the mouth, etc., are meaningful to humans and can be readily extracted from the mouth images. Alternatively, model-based visual features are typically obtained in conjunction with a parametric or statistical facial feature extraction algorithm. With model-based approaches, the model parameters are directly used as visual features [5, 12]. Examples of



Lip Movement Recognition. **Figure 3** Various visual facial features divided into appearance- and shape-based features.

statistical models are active shape and active appearance models [13].

The combination of appearance- and shape-based visual features has also been utilized in expectation of improving the performance of the recognition system, since they contain respectively low- and high-level information about the person's lip movements. Appearance- and shape-based features are usually just concatenated, or a single model of face shape and appearance is created [13]. The dynamics of the changes of visual features are usually captured by augmenting the visual feature vector by its first- and second-order time derivatives, computed over a short temporal window centered at the current video frame.

Finally, in the classification phase, a number of classifiers can be used to model prior knowledge of how the visual features are generated by each speaker. They are usually statistical in nature and utilize ANNs, SVMs, GMMs, HMMs, etc. The parameters of the prior model are estimated during training. During testing, based on the trained model the posterior probability is maximized, and identification/verification decision is made.

Summary

Lip movement recognition and audio–visual-dynamic speaker recognition are speaker recognition technologies that are user-friendly, low-cost, and resilient to spoofing. There are many biometric applications, such as, sport venue entrance check, access to desktop, building access, etc., in which it is very important to use unobtrusive methods for extracting biometric features, thus enabling natural person recognition and reducing inconvenience. Low cost of audio and video biometric sensors and the ease of acquiring audio and video signals (even without assistance from the client), makes biometric technology more socially acceptable and accelerates its integration into every day life.

Related Entries

- ▶ Face Recognition
- ▶ Face Tracking
- ▶ Multibiometrics
- ▶ Multibiometrics and Data Fusion
- ▶ Multimodal Systems
- ▶ Speaker Matching
- ▶ Session Effects on Speaker Modeling
- ▶ Speaker Recognition, Overview
- ▶ Speech Analysis
- ▶ Speech Production
- ▶ Spoofing

References

1. Chen, T., Rao, R.R.: Audio-visual integration in multimodal communication. Proc. IEEE **86**(5), 837–852 (1998)
2. Aleksic, P.S., Potamianos, G., Katsaggelos, A.K.: Exploiting visual information in automatic speech processing. In: Bovik, A.L. (ed.) Handbook of Image and Video Processing. Academic, London (2005)
3. Aleksic, P.S., Katsaggelos, A.K.: Speech-to-video synthesis using MPEG-4 compliant visual features. IEEE Trans CSVT, Special Issue on Audio and Video Analysis for Multimedia Interactive Services, pp. 682–692, May (2004)
4. Summerfield, A.Q.: Some preliminaries to a comprehensive account of audio-visual speech perception. In: Campbell, R., Dodd, B. (eds.) Hearing by Eye: The Psychology of Lip-Reading, pp. 3–51. Lawrence Erlbaum, London, United Kingdom (1987)
5. Aleksic, P.S. Katsaggelos, A.K.: Audio-visual biometrics. IEEE Proc **94**(11), 2025–2044 (2006)
6. Chaudhari, U.V., Ramaswamy, G.N., Potamianos, G., Neti, C.: Audio-visual speaker recognition using time-varying stream reliability prediction. IEEE Proc. Int. Conf. Acoustics Speech Signal Process. (Hong Kong, China) **5**, V-712–15 (2003)
7. Hjelmas, E., Low, B.K.: Face detection: A survey. Computer Vision. Image Understand. **83**(3), 236–274 (2001)
8. Hennecke, M.E., Stork, D.G., Prasad, K.V.: Visionary speech: Looking ahead to practical speechreading systems. In: Stork, D.G., Hennecke, M.E. (eds.) Speechreading by Humans and Machines, pp. 331–349. Springer, Berlin (1996)
9. Aleksic, P.S., Katsaggelos, A.K.: Comparison of low- and high-level visual features for audio-visual continuous automatic speech recognition. IEEE Proc. Int. Conf. Acoustics Speech Signal Process. (Montreal, Canada) **5**, 917–920 (2004)
10. Potamianos, G., Graf, H.P., Cosatto, E.: An image transform approach for HMM based automatic lipreading. Paper presented at the Proceedings of the International Conference on Image Processing, vol. 1, pp. 173–177. Chicago, IL, 4–7 Oct. 1998
11. Wark, T., Sridharan, S., Chandran, V.: Robust speaker verification via fusion of speech and lip modalities. Proc. Int. Conf. Acoustics Speech Signal Process. Phoenix **6**, 3061–3064 (1999)
12. Aleksic, P.S., Katsaggelos, A.K.: An audio-visual person identification and verification system using FAPs as visual features. Paper presented at the Proceedings of Works. Multimedia User Authentication, pp. 80–84. Santa Barbara, CA (2003)
13. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. Paper presented at the Proceedings of European Conference on Computer Vision, pp. 484–498. Freiburg, Germany (1998)

Lip-Radiation Effect

The lip-radiation effect corresponds to changing the volume velocity waveform at the lips to a speech pressure signal in a free field at a certain distance from the speaker.

► Speech Production

Liquid Crystal Displays (LCD)

Liquid crystal display (LCD) is a digital display that uses liquid crystal cells whose reflectivity varies according to the voltage applied to them. Liquid crystal is a substance that flows like liquid but maintains some of the ordered structure characteristic of the crystals.

► Digitizing Tablet

Liveness Detection

► Fingerprint Fake Detection
► Liveness Iris

Liveness Assurance in Face Authentication

Michael Wagner, Girija Chetty
School of Information Sciences, University of Canberra, Australia

Synonyms

Face authentication; Face recognition; Face verification

Definition

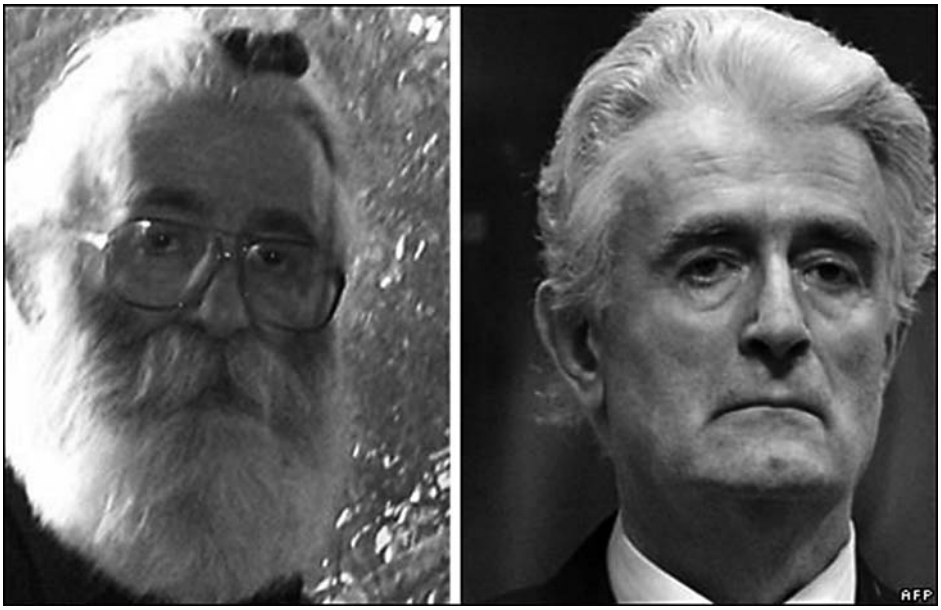
The process of verifying whether the face image presented to an authentication system is real (i.e., alive), or whether it is reproduced or synthetic, and thus fraudulent. When a face authentication system is to recognize the face of a person by means of an electronic camera and associated image recognition software, it is important to be sure that the person seeking the authentication actually presents his or her face to the camera at the time and place of the authentication request. The face is presented live as on live television as distinct from a movie or cartoon programme. In contrast, an impostor could try to present a photograph, or a video recording, of a legitimate client to the camera in order to be falsely authenticated by the system as that client. That kind of threat to authentication systems is known generally as ► replay attack. In turn, liveness assurance uses a range of measures to reduce the vulnerability of face authentication systems to the threats of replay attack.

Introduction

The primary design objective for a face recognition system is its ability to distinguish, as clearly as possible, between different persons on the basis of their facial images. As is described in detail in sections ► Face Recognition, Overview, a good face recognition system utilizes a suitable feature set, employs sophisticated pattern recognition algorithms and sets decision thresholds appropriate for the specific application context. Nevertheless, current face recognition technology is vulnerable on several fronts: on one hand, different persons like twins, especially identical twins, other siblings, parents and children can have quite similar facial appearance, while on the other hand the same person can appear quite dissimilar at different times owing to facial hair and hairstyle, make-up, cosmetic surgery, eye glasses, and even just due to their physical or emotional state. Figure 1 shows an example of the faces of two identical twins being almost indistinguishable and Fig. 2, in contrast, shows the large difference between two images of the same person who changed his facial appearance drastically. Face recognition also has robustness issues unless environment variables such as lighting of the face and pose with respect to the camera are controlled meticulously.

In addition, face authentication systems are vulnerable to impostors who present a photograph of a legitimate client to the system camera and may be falsely accepted as that client by the system [1]. Generally, such an attack on a biometric authentication

Liveness Assurance in Face Authentication. **Figure 1** Similarity of the facial images of two different persons (downloaded from http://www.mary-kateandashley.com).



Liveness Assurance in Face Authentication. **Figure 2** Dissimilarity of two facial images of the same person (downloaded from http://news.bbc.co.uk).

system is known as a replay attack. Replay attacks can be carried out by presenting a printed photograph to the system camera or by holding a computer screen showing a photo or video recording in front of the camera. However, replay attacks are also possible by injecting a suitable recorded signal or data file at other points within the authentication system. All replay attacks have in common that, at the time of the authentication they play back to the system a signal that was recorded from the client at an earlier time.

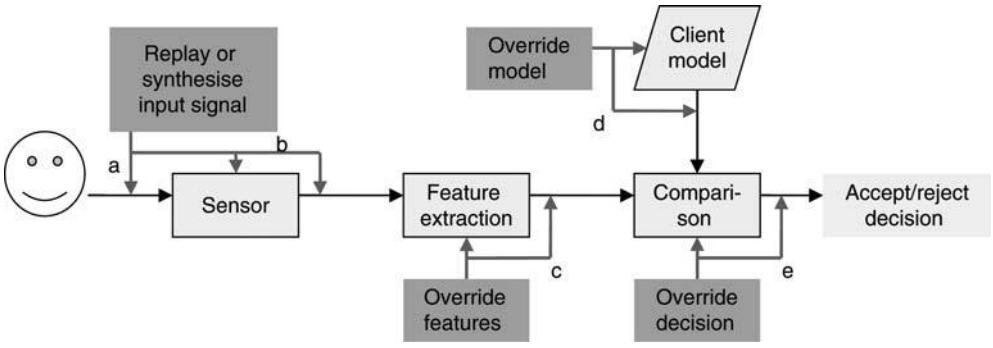
Closely related to the replay attack is another form of attack on a face authentication system, namely the ► synthesis attack. A synthesis attack does not use a pre-recorded signal, such as a photograph or video of a client, directly. Instead, it uses known client data to build a client model, for example a three-dimensional shape and texture model of the client's head. From such a model, entirely artificial photographs or video sequences with or without speech sounds can be synthesized, which can closely resemble the actual client.

Replay Attack

The different points at which a face recognition system is vulnerable to replay attacks are shown in Fig. 3. An attacker can present a photograph or play back a video of the face of a true client to the sensor, or electronic camera, of the authentication system as shown in Fig. 3(a). This point is the most vulnerable in the authentication system because in a fully automated system, the possibility of presenting a photograph is always available to an attacker unless the physical space in front of the camera is supervised by a human observer or by a second biometric modality in addition to the facial image camera. If an attacker can gain access to the inside of the camera or to the connection between the camera and the back end of the system, as shown as in Fig. 3(b), the attacker does not need to "show" a physical photograph or video to the camera, but can inject a suitable electronic signal that corresponds to the facial image of the client into the system directly.

Since a face authentication system will invariably be implemented as software running on a computer or network of computers, such a system would be open to the same threats as any other software, particularly if it is connected to the Internet. The vulnerabilities of computer systems to a range of threats, including viruses, worms, Trojan Horses, or even more simply, the disclosure or easy guess of passwords, are well known and any biometric system is subject to those same threats. Accordingly, if an attacker can gain access to the face authentication system at or beyond the feature extraction stage, as shown as in Fig. 3c, the attacker can bypass the input of a facial image altogether and present the system with the fake features of a client face. An attacker who is able to access the stored client models of the system, shown in Fig. 3d, will achieve the ultimate identity theft by replacing the model of a real client with the model of an impostor. This will have the effect that forthwith the impostor will be falsely accepted by the system as the client since the impostors face will, of course, now be compared with the impostor's own facial model, which has been substituted for the model of the real client.

The ultimate success for an attack of the face authentication system lies in the attacker being able to access the Comparison Module of the system, as shown in Fig. 3e, since a breach of the system at that point will enable the attacker to override the system with his own accept or reject decision irrespective of the face shown to the camera or the client model that the face is compared with.



Liveness Assurance in Face Authentication. Figure 3 Potential points of vulnerability of a face authentication system: (**a**) replay or synthesize the client facial image into the input sensor; (**b**) insert the replayed or synthesized client facial image into vulnerable system-internal points; (**c**) override detected features at vulnerable system-internal points; (**d**) override the client at vulnerable system-internal points; (**e**) override the accept/reject decision at vulnerable system-internal points.

Liveness Assurance for Face Authentication: Visual Sensors Only

Additional Infrared or Ultraviolet Sensors

Depending on the nature of the replay attack, different methods of liveness assurance can be used. A still photograph or a video presented to the system camera as a paper print or on a computer screen will always reflect the spectral sensitivity of the recording device. Therefore, a system camera, which has a different spectral sensitivity from that of an ordinary camera, for example extending into either the infrared or ultraviolet range of the spectrum, is able to distinguish a live face from a photo or video recorded with an ordinary camera. An infrared or ultraviolet camera can also be employed as a secondary input device in addition to an ordinary-spectrum camera [2]. Such secondary sensors, which, for example, could show the temperature profile of the face or the vein pattern underneath the skin, are excellent liveness detectors, provided that the training of the client models is undertaken with the same sensor arrangement that is later used for client authentication. However, the disadvantage of such a sensor arrangement is that infrared and ultraviolet sensors are expensive. Moreover, such sensors cannot be used where the authentication system has to rely on ordinary cameras such as webcams or mobile phone cameras in a distributed authentication system.

Detection of 3D Head Movement

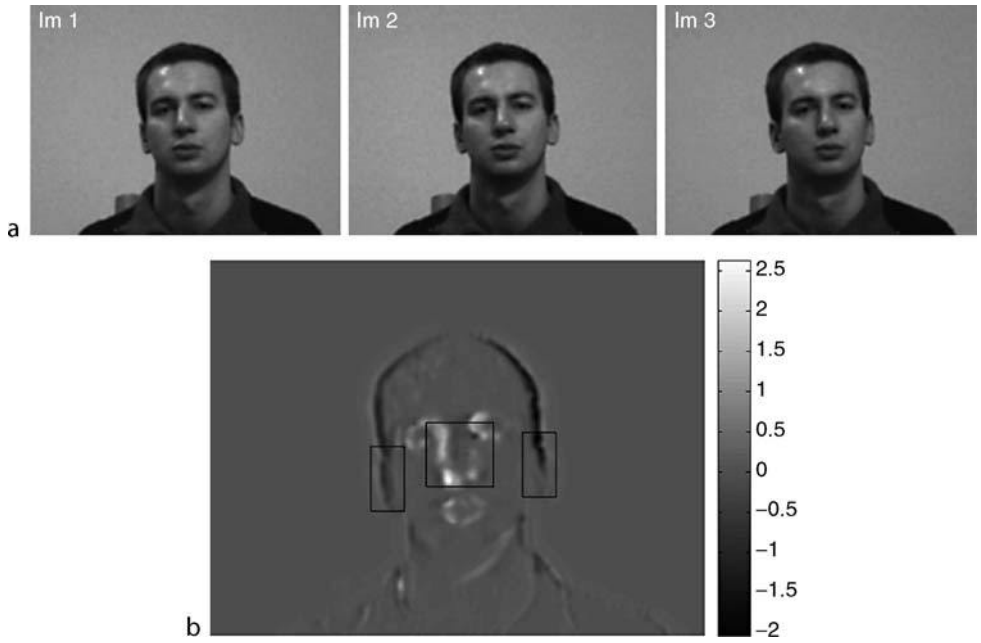
Another method of distinguishing a live face from a photograph or video is to ascertain that the face as it is presented to the camera moves in a manner consistent with the three-dimensional shape of the human head. A rotation of a real human head in front of a camera will reveal parts of the head that were obscured prior to the rotation, while at the same time obscuring other parts that were previously visible. This effect distinguishes the rotation of a real human head from the rotation of a photograph or image on a computer screen. More generally, the positions of facial "landmarks," such as pupils, nose tip, or mouth corners of a three-dimensional head – and hence the distances between such landmarks – will follow the rules of three-dimensional trigonometry and as such can be distinguished clearly from rotations of a two-dimensional photograph or computer screen.

The detection of a three-dimensional head can be achieved either by utilising a stereo camera – or several cameras looking at the head from different directions – or by taking a sequence of images of the moving head through a single camera. In the first case, the presentation of a two-dimensional photograph or computer screen is immediately obvious, while in the second case the system can either make use of inadvertent small rotations of the client's head or explicitly ask clients to rotate their heads in a prescribed manner.

An example of a system, which uses an image sequence collected by a single camera in order to detect three-dimensional head movements, is described by Kollreider et al. [3]. The system is based on the observation that the two-dimensional image of a head rotating around its vertical axis shows significant lateral movement in the centre of the face while the ears, forehead, and chin move mostly in directions perpendicular to the projection. Optical flow estimation ▶ Face Recognition, 3D-Based and face part detection are used to measure and compare the movements of the nose and the ears, respectively, across an image sequence of a rotating head. If the lateral movement of the nose over the time span of the image sequence is larger than the lateral movement of the ears, it is assumed that a real head is rotating in three dimensions rather than a two-dimensional image being turned in front of the camera. By an appropriate threshold on the difference of the horizontal pixel velocities between the nose region and the ear regions, video sequences of a three-dimensional rotating human head are distinguished from those of a two-dimensional rotated photograph. Figure 4a shows three frames of a head rotation sequence, and Fig. 4(b) shows the corresponding optical flow diagram.

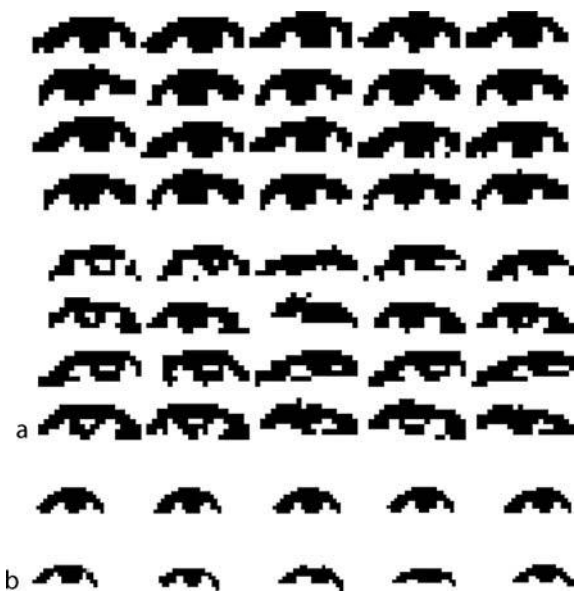
Detection of Facial Micro-Movement

Another possibility to distinguish a live face from a photograph is based on the assumption that an image sequence of a live face will invariably show some variation of facial features. This is obviously the case when the person is speaking and there is facial variation, mainly in the mouth region, which corresponds to



Liveness Assurance in Face Authentication. Figure 4 (**a**) Image sequence of rotating head; (**b**) horizontal optical flow magnitude showing higher pixel velocities (white) for the central area and lower pixel velocities for the peripheral areas of the face. ([3] © 2005 IEEE.)

the speech sounds being produced. Under an assumption that a person will always exhibit some eye movement over time, it is possible to distinguish a live face from a photograph by comparing the eye regions between consecutive frames of a video sequence. A system, which utilizes inadvertent eye movements to distinguish between a live face and a photograph is proposed by Jee et al. [4]. The system uses five sequential face images, then detects the centre points of both eyes in order to extract the two eye regions. For each of the eye regions, the 20×10 pixels of that region are 1-bit quantized to be either black or white and Hamming distances are calculated between the five consecutive images of each eye region. Figure 5(a) shows the sequences of five black and white frames for the eye regions of still photographs, while Fig. 5(b) shows frame sequences of the eye regions of live faces. The figure clearly shows a larger variation of the eye regions of the live faces than of the still photos. According to Jee et al. [4], live images can be distinguished from photographs because the average Hamming distance between the five images of a sequence is always larger for a live face than for a still photograph presented to the camera.



Liveness Assurance in Face Authentication. Figure 5 (**a**) Five consecutive video frames of eye regions from still photos; (**b**) five consecutive video frames of eye regions from live faces. Pixels are 1-bit quantized to black or white ([4] Courtesy World Academy of Science, Engineering and Technology.)

Challenge-Response Paradigm

In addition, it is possible for the authentication system to issue a "challenge" to the persons seeking authentication by asking them to perform some specific prescribed movement, for example "tilt head to the left" or "blink with your right eye" as proposed in the Facial Liveness Assessment System [2]. Such system requests are akin to the prompted-text paradigm in speaker recognition (► Liveness Assurance in Voice Authentication). They provide a good defence against replay attack, but on the system side it is necessary to design and implement an automatic mechanism, which is able to reliably confirm the correctness of the client's response to the system prompt.

Vulnerability to Replayed Video Recordings

The above liveness assurance methods provide protection against forms of replay attack that present a recorded image of a client's face to the system camera. If the attacker uses printed photographs or photographs displayed on a computer screen, particularly on the display of an easily portable notebook computer, all of these methods provide good distinction between a replay attack and a client face that is presented live to the camera. However, an attacker who is able to present a client photograph to the system on a notebook screen, is also likely to be capable of replaying a recorded video sequence on such a notebook. In this case, the single-camera 3D detection method will fail to detect a replay attack because the recorded video sequence has the same three-dimensional rotation characteristic as a human head rotated live in front of the system camera. Similarly, the detection of micro-movements would fail because the video recording contains the same facial micro-movements of the lip and eye regions as a human face presented live to the camera. The only system architecture that is capable – without the presence of a human supervisor – of distinguishing between a two-dimensional video presentation and a three-dimensional live presentation of a human face, is one that has a three-dimensional sensor arrangement with a set of cameras surrounding the head of the client and/or obtaining a wide-angle view of the scene, and as such is able to "look behind" a two-dimensional printed photograph or notebook computer held before the cameras.

Multimodal Liveness Assurance

While it is feasible to deceive a single-camera system by replaying a video recording on a notebook, held in front of the camera, it is far more difficult to use the same notebook to deceive an acoustic speaker recognition system by replaying a sound recording through the notebook's in-built speakers or another small loudspeaker. From the point of view of attackers, there are several obstacles: firstly they must not be detected holding a computer screen in front of the camera; secondly they must provide a high-quality loudspeaker, which is usually bulky and not normally found in notebook computers; and thirdly they have to play back a recorded video with perfectly synchronous facial images and speech sounds. Therefore, a multimodal approach to liveness assurance has been proposed, which combines the recognition of a client's face with the recognition of the client's voice [5]. In a combined face-voice authentication system it is possible to verify not just that there are some – random – micro-movements in the lip area of the face, but that those lip movements correspond precisely to the speech sounds that are heard simultaneously by the system microphone. For example, the labial consonant /p/ in "Paris" would correspond to a closing followed by an opening of the lips, while the rounded vowels /u/ in "Toulouse" would correspond to a rounded lip configuration.

More generally, the assurance of liveness in a bimodal face-voice authentication system is based on the fact that the articulator movements, mainly of the lips, but also of the tip of the tongue, jaw, and cheeks are mostly observable and correspond closely to the particular speech sounds produced. Therefore it is possible when observing a bimodal audio-video signal of the speaking face to ascertain whether the facial dynamics and the sequence of speech sounds are mutually compatible and synchronous. Human observers are finely tuned to the synchrony of acoustic and visual signal and it is quite disconcerting when one or the other is delayed or there is no apparent match, for example, with an out-of-sync television signal or with a static facial image when the speaker is heard saying

something, but the lips are not seen to be moving. In the field of audiovisual speech recognition, the term "viseme" has been coined as the visual counterpart of the "phoneme," which denotes a single speech sound ▶ Speaker Recognition, Overview. This is illustrated in Fig. 6, which shows a set of corresponding phonemes and visemes of English. The visemes /m/, /u/, and /d/ (as in the word "mood"), for example, first show the speaker's lips spread and closed (for /m/), then protruded and rounded (for /u/), and finally spread and slightly open (for /d/). It is therefore possible to detect whether corresponding sequences of visemes and phonemes of an utterance are observed in a bimodal audio-video signal and whether the observed viseme and phoneme sequences are synchronous.

Vis	Phoneme		Example	Mouth
	MPEG-4	BEEP eq.		
0	none (#)	sil	N/A	
1	p,b,m	p,b,m	put,bed,me	
2	f,v	f,v	far,voice	
3	T,D	th,dh	think,the	
4	t,d	t,d	tip,do	
5	k,g	k,g	cold,gap	
6	tS,dZ,S	ch,jh,sh	chair, join,she	
7	s,z	s,z	sir,zeal	
8	n,l	n,l	not,loud	
9	r	r	red	
10	A	aa,ae,ah	car,and	
11	e	eh,ax	bed,the	
12	I	ih,iy	in,me	
13	U	aw,uh,uw	loud,book,do	
14	-	er	sir	
15	-	w	with	
16	O	oh,ow,oy	top,cold,voice	

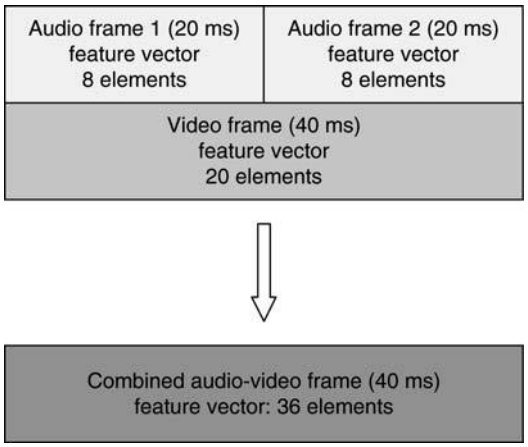
Liveness Assurance in Face Authentication. **Figure 6** Visemes and their corresponding phones.

In order for the synchrony of the audio and video streams to be ascertained, the two modalities must be combined appropriately. Multimodal authentication systems employ different paradigms to combine, or "fuse," information from the different modalities. Modality fusion can happen at different stages of the authentication process. Fusing the features of the different channels immediately after the feature extraction phase is known as "feature fusion" or "early fusion." In this paradigm, all comparisons between the unknown sample and the client model as well as the decision making are based on the combined feature vectors. The other possibility is to fuse information from the two modalities after independent comparisons have been made for each modality.

For liveness assurance by means of bimodal face-voice authentication, it is necessary to apply an early fusion stratagem, i.e., to fuse the two modalities at the feature level [6]. If the two modalities were to be fused late, i.e., at the score or decision level, analysis of the video of the speaking face would yield one decision on the speaker's identity and analysis of the audio of the utterance would yield another decision on the speaker's identity. The two processes would run independently of each other with no connection between them that would allow the checking for the correspondence and synchrony of visemes and phonemes [7].

Therefore, the features that are extracted from the audio signal on a frame-by-frame basis – usually at an audio frame rate of about 40–100 frames per second – must be combined with the features that are extracted from the video signal – usually at the video frame rate of 25 or 30 frames per second. An example of how the differing frame rates for the audio and video signals can be accommodated is shown in Fig. 7, where the audio frame rate is 50 frames per second, the video frame rate is 25 frames per second, and the combined audiovisual feature vector comprises the audio feature vectors of two consecutive audio frames, combined with the single video vector of the synchronous video frame.

The combined audiovisual feature vectors will then reveal whether the audio and video streams are synchronous, for example when the combined audiovisual feature vectors contain the sequence of visemes /m/, /u/, and /d/ and likewise the sequence of phonemes /m/, /u/, and /d/. In contrast, if one of the combined audiovisual feature vectors were to contain the visual information for the viseme /m/ and at the same time the



Liveness Assurance in Face Authentication. Figure 7 Feature fusion of two consecutive 20 ms audio feature vectors with the corresponding 40 ms video feature vector. Before fusion, the audio vectors have been reduced to 8 dimensions each, and the video vector has been reduced to 20 dimensions. The combined feature vector has 36 dimensions [5].

audio information for the phoneme /u/, that combined feature vector would indicate that the audio and video streams do not represent a corresponding synchronous representation of any speech sound.

The proper sequencing of visemes and phonemes is usually ascertained by representing the audiovisual speech by Hidden Markov Models (HMM), which establish the likelihoods of the different combined audiovisual vectors and their sequences over time [8]. It is therefore possible to ascertain whether the audio and video components of a combined audio-video stream represent a likely live utterance. Therefore, an attacker who attempts to impersonate a target speaker by means of a recorded speech utterance and a still photograph of the target speaker will be thwarted because the system will recognize the failure of the face to form the corresponding visemes that should be observed synchronously with the phonemes of the utterance. Similarly, such a system will thwart an attack by an audiovisual speech synthesis system, unless the synthesizer can generate the synthetic face and the synthetic voice in nearly perfect synchrony.

The combination of face authentication with voice authentication has a number of advantages beyond the assurance of the liveness of the face-voice samples. Firstly, the method is well supported by telecommunication devices, which are increasingly equipped with image and sound sensors that are capable of delivering facial images and voice samples suitable for remote client authentication. Secondly, the combination of two largely – but not completely! – independent biometrics has an advantage in terms of error rates compared with either modality employed singly. And thirdly, the utilisation of combined image and sound signals has a distinct advantage in robustness when the environmental conditions are adverse to either the face recognition system or the speaker recognition system, for example when lighting conditions are not conducive to successful face recognition, or when a passing train makes speaker recognition all but impossible.

Related Entries

- ▶ Face Recognition, 3D Based
- ▶ Face Recognition, Overview
- ▶ Liveness: Detection Voice
- ▶ Optical Flow
- ▶ Speaker Recognition, Overview

References

1. Schuckers, S.A.C.: Spoofing and anti-spoofing measures, Information Security Technical Report. **7**(4), 56–62 (2002)
2. Facial liveness assessment system, Int. Patent WO/2005/008566, http://www.wipo.int (27.1.2005)
3. Kollreider, K., Fronthaler, H., Bigun, J.: Evaluating liveness by face images and the structure tensor. In: Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies, 17–18 Oct 2005. IEEE (2005)
4. Jee, H.-K., Jung, S.-U., Yoo, J.-H.: Liveness detection for embedded face recognition system. In: Proceedings of World Academy of Science, Engineering and Technology, vol. 18, Dec 2006, ISSN 1307–6884 (2006)
5. Chetty, G., Wagner, M.: "Liveness" verification in audio-video authentication. In: Proceedings of the International Conference on Spoken Language Processing, ICSLP-2004, Jeju, Korea, 4–7 Oct 2004, vol. III, pp 2509–2512 (2004)
6. Chetty, G., Wagner, M.: Investigating feature level fusion for checking liveness in face voice authentication. In: Proceedings of the Eighth International Symposium on Signal Processing and Its Applications, Sydney, 28–31 Aug 2005 (2005)
7. Bredin, H., Chollet, G.: Audiovisual speech synchrony measure: application to biometrics. EURASIP J. Adv. Signal Process. 2007 (1), 1–11 (2007)
8. Chetty, G., Wagner M.: Speaking faces for face-voice speaker identity verification. In: Proceedings of Interspeech-2006 – International Conference on Spoken Language Processing, Paper Mon3A1O-6, Pittsburgh. International Speech Communication Association (2006)

Liveness Assurance in Voice Authentication

Michael Wagner
School of Information Sciences, University of Canberra, Australia

Synonyms

One-to-One Speaker recognition; Speaker verification; Voice authentication; Voice verification

Definition

The process of verifying whether the voice sample presented to an authentication system is real (i.e., alive), or whether it is replayed or synthetic, and thus fraudulent. When authentication through a ► voice authentication system is requested, it is important to be sure that the person seeking the authentication actually provides the required voice sample at the time and place of the authentication request. The voice is presented live like that of a radio presenter during a live broadcast as distinct from a recorded audio tape. In contrast, an impostor who seeks authentication fraudulently could try to play an audio recording of a legitimate client or synthesized speech that is manufactured to resemble the speech of a legitimate client. Such threats to the system are known as ► replay attack and ► synthesis attack, respectively. ► Liveness assurance uses a range of measures to reduce the vulnerability of a voice authentication system to the threats of replay and synthesis attack.

Introduction

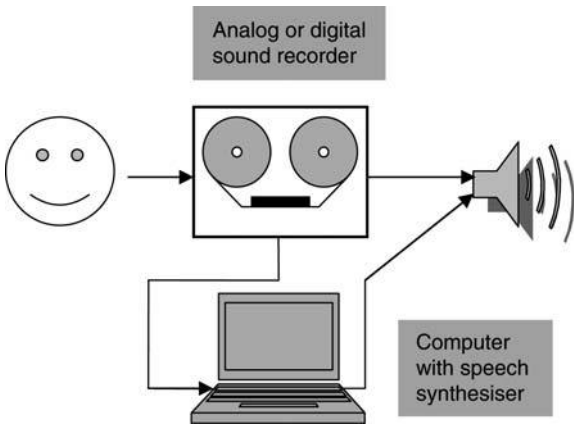
The security of a voice authentication system depends on several factors ► Voice Authentication. Primarily it is important that the system is capable of distinguishing people by their voices, so that a clients who are enrolled in, say, a telephone banking system are admitted to their account reliably, while an "impostor" who attempts to access the same account is rejected equally reliably. A good voice authentication system will thwart an impostor irrespective of whether the access to the other person's account is inadvertent or deliberate and irrespective of whether the impostors use their natural voice or try to improve their chances by mimicking the voice of the client.

However, one vulnerability common to all voice authentication systems is the possibility that attackers, instead of speaking to the system directly and with their own voice, fraudulently use the recorded voice of a true client in order to be admitted by the system. In principle, such a "replay attack" can be carried out by means of any sound recording device, analogue or digital, through which the recorded voice of the client is played back to the system, say, to a microphone at a system access point or remotely into a telephone handset connected to the authentication system. The security issue in this case is that the voice used for authentication is not the "live" voice of the person, who is seeking access to the system, at the time and place of the access request.

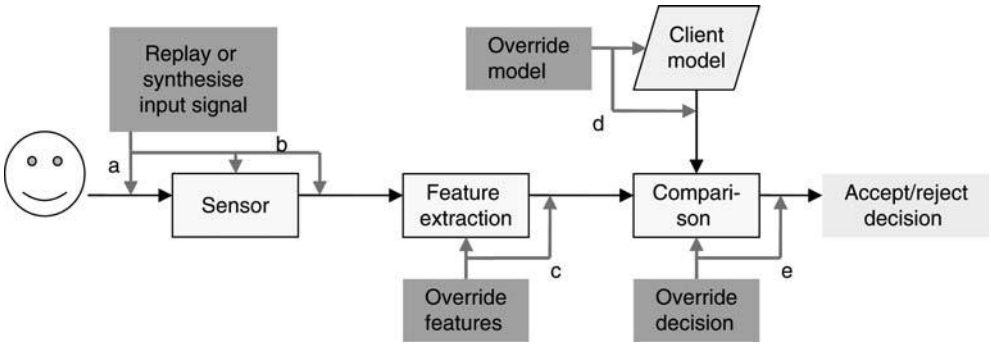
A technically sophisticated attacker may also use suitable computer hardware and software to create a simile of the client's voice by means of speech synthesis without having to record specific voice samples of the client. Such an attack will be referred to as a "synthesis attack" in the following. Figure 1 shows how replayed or synthesized voice signals can be substituted for the live voice of a client at the sensor input of the authentication system.

Replay Attack

Since voice authentication is always implemented within the context of a computer system, it is



Liveness Assurance in Voice Authentication. Figure 1 Prerecording of client voice either for later replay or for generating a client model, which can be used later to synthesize the client's voice.



Liveness Assurance in Voice Authentication. Figure 2 Potential points of vulnerability of a voice biometric authentication system: (**a**) replay or synthesize the client voice into the input sensor; (**b**) insert the replayed or synthesized client voice into vulnerable system-internal points; (**c**) override detected features at vulnerable system-internal points; (**d**) override the client at vulnerable system-internal points; (**e**) override the accept/reject decision at vulnerable system-internal points.

important to consider the vulnerabilities of the entire system generally (▶ Security and Liveness, Overview). Figure 2 shows the structure of a typical voice authentication system. During the enrolment or training phase, the client's voice is captured by the microphone, salient features are extracted from the speech signals and finally a statistical "client model" or template is computed, which represents the client-specific voice characteristics according to the speech data collected during enrolment. During the operational or testing phase when the system needs to decide whether a speech sample belongs to the client, the signal is also captured by the sensor and features are extracted in the same way as they are in the enrolment phase. Then, the features of the unknown speech sample are compared statistically with the model of the client that was established during enrolment. Depending on how close the unknown sample is to the client model, the system will issue either an "accept" or a "reject" decision: the person providing the voice sample is either authenticated or considered an impostor.

Figure 2 shows various ways in which attackers could manipulate the outcome of the authentication, if any of the software or hardware components of an insecure computer system could be accessed. If it were possible, for example, to manipulate the database of client models, attackers could potentially replace the voice model of a client with their own voice model and subsequently gain fraudulent access to the system by having substituted their own identity for that of the client. Or, even more simply, if it were possible to manipulate the decision module of the system, an attacker could essentially bypass the entire authentication process and manufacture an "accept" decision of the system without having provided any matching voice data. Such considerations fall into the domain of the system engineer who needs to ensure, much as with any other secure system, that there are no bugs, trap doors, or entry points for Trojan Horses, which could allow an attacker to manipulate or bypass the authentication mechanisms of the system. Since such vulnerabilities are not specific to voice authentication systems, they are not dealt with in this essay (▶ Biometric Vulnerabilities: Overview).

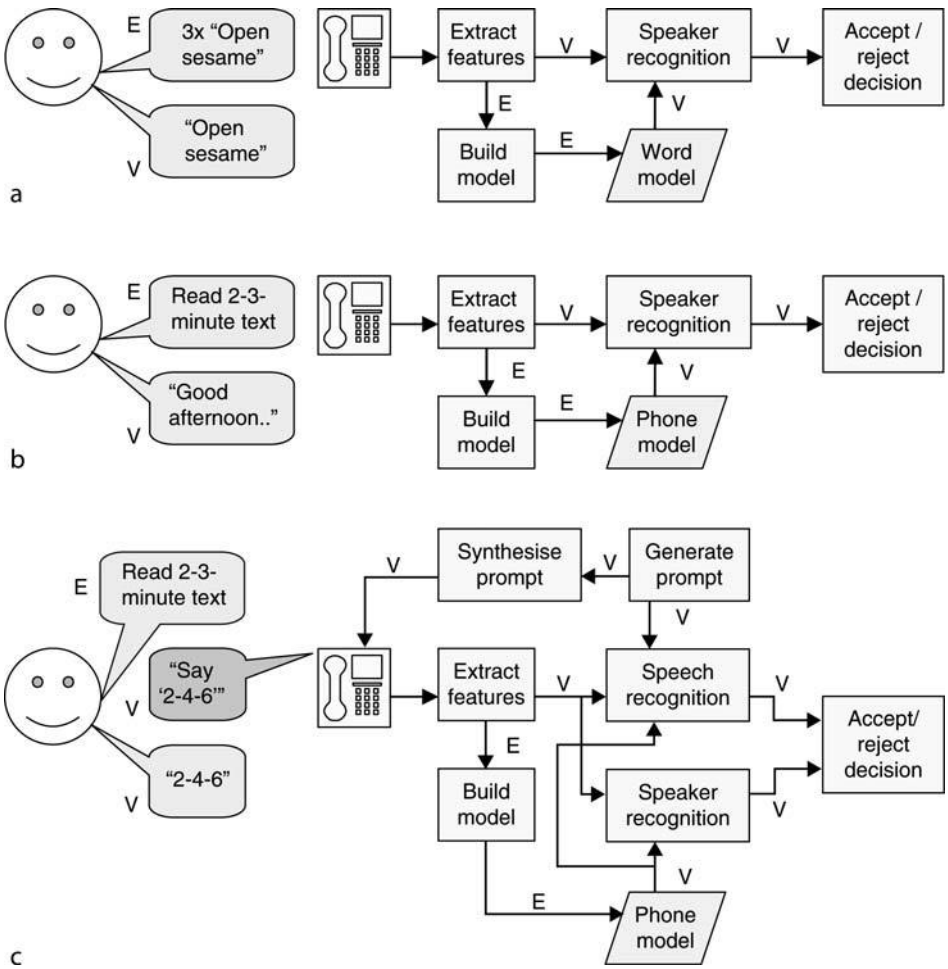
The remainder of this article discusses how a secure voice authentication system can provide the assurance that the voice used for an access request to the system is "live" at the time and place of the access request and is neither a playback of a voice recording nor a synthesized simile of a client voice. Hence, liveness assurance is an essential aspect of the security of any voice authentication system.

Liveness Assurance for Different Authentication Protocols

Voice authentication systems operate under different protocols and assurance of liveness is affected differently by the various authentication protocols. The three main protocols used for voice authentication are text-dependent speaker verification, text-independent speaker verification, and text-prompted speaker

verification, as shown in Fig. 3. The earliest authentication protocol were text-dependent [1]. In this protocol, the client uses a fixed authentication phrase, which is repeated several times during enrolment. The repetitions are necessary so that the system "learns" about the range of pronunciation of the authentication phrase by the client. Generally, speaker verification works best if the natural variation of a speaker's voice is well captured during enrolment. Hence, ideally, enrolment should be distributed over several recording sessions that may be spread over several days or even weeks. The same phrase, for example, a sequence of digits ("three-five-seven-nine"), or a password or passphrase ("Open Sesame") is then used again by the client during the operational phase in order to be authenticated by the system.

Text-dependent systems have the advantage that the client model only needs to represent the acoustic information related to the relatively few speech sounds of the passphrase. Enrolment, therefore, is shorter and quicker than for other protocols, which typically require the representation of the entire collection of speech sounds that the client could possibly produce. However, the text-dependent protocol has the distinct disadvantage that clients will repeat the same phrase every time while using the system. Consequently, there



Liveness Assurance in Voice Authentication. Figure 3 (**a**) Text-dependent voice authentication: (E) at enrolment the client repeats the authentication phrase several times; (V) for verification the client speaks the same authentication phrase. (**b**) Text-independent voice authentication: (E) at enrolment the client reads a 2–3-min phonetically-rich text; (V) for verification any utterance can be used by the client. (**c**) Text-prompted voice authentication: (E) at enrolment the client reads a 2–3-min phonetically rich text; (V) for verification the client is prompted to say a given phrase, which is verified both for the correct content and for the client's voice characteristics.

may be ample opportunity for an attacker, especially if the system microphone is situated in a public area, to plan and carry out a surreptitious recording of the passphrase, uttered by the client, and to replay the recorded client passphrase fraudulently in order to be authorized by the system.

In contrast, text-independent voice authentication systems [2] will authenticate a client – and reject an impostor – irrespective of any particular utterance used during enrolment. Client enrolment for text-independent systems invariably takes longer than enrolment for a text-dependent system and usually involves a judiciously designed enrolment text, which contains all, or at least most, of the speech sounds of the language. This will ensure that the client models, which are constructed from the enrolment speech data, will represent to the largest extent possible the idiosyncrasies of the client when an arbitrary sentence or other utterance is provided for authentication later. Text-independent protocols offer the advantage that authentication can be carried out without the need for a particular passphrase, for example, as part of an ordinary interaction between a client and a customer-service agent or automated call centre agent, as shown in this fictitious dialog:

Client phones XYZ Bank.

Agent: Good morning, this is XYZ Bank. How can I help you?

Client: I would like to enquire about my account balance.

Agent: What is your account number?

Client: It's 123-4567-89

Agent: Good morning, Ms Applegate, the balance of your account number 123-4567-89 is $765.43. Is there anything else. . .?

The example shows a system, which combines speech recognition with voice authentication. The speech recognizer understands what the customer wants to know and recognizes the account number, while the authentication systems uses the text-independent protocol to ascertain the identity of the client from the first two responses the client gives over the telephone. These responses would not normally have been encountered by the system during enrolment, but the coverage of the different speech sounds during enrolment would be sufficient for the authentication system to verify the client from the new phrases. The text-independent protocol offers an attacker the opportunity to record *any* client utterances either in the context of the client using the authentication system or elsewhere, and to replay the recorded client speech in order to fraudulently achieve authentication by the system.

A more secure variant of the text-independent protocol is the text-prompted protocol [3]. Enrolment under this protocol is similar to the text-independent protocol in that it aims to achieve a comprehensive coverage of the different possible speech sounds of a client so that later on any utterance can be used for client authentication. However, during authentication the text-prompted protocol asks the user to say a specific, randomly chosen phrase, for example, by prompting the user "please say the number sequence 'two-four-six'". When the client repeats the prompted text, the system uses automatic speech recognition to verify that the client has spoken the correct phrase. At the same time it verifies the client's voice by means of the text-independent voice authentication paradigm. The text-prompted protocol makes a replay attack more difficult because an attacker would be unlikely to have all possible prompted texts from the client recorded in advance. However, such an attack would still be feasible for an attacker with a digital playback device that could construct the prompted text at the press of a button. For example, an attacker who has managed surreptitiously to record the ten digits "zero" to "nine" from a client – either on a single occasion or on several separate occasions – could store those recorded digits on a notebook computer and then combine them to any prompted digit sequence by simply pressing buttons on the computer.

Synthesis Attack

Even a text-prompted authentication system is vulnerable to an attacker who uses a text-to-speech (TTS) synthesizer. A TTS system allows a user to input any desired text, for example, by means of a computer keyboard, and to have that text rendered automatically into a spoken utterance and output through a loudspeaker or another analog or digital output channel. The basic principle is that an attacker would program a TTS synthesizer in such a way that it produces similar speech patterns as the target speaker. If that is achieved, the attacker would only need to type the text that is required or prompted by the authentication system in order for the TTS synthesizer to play the equivalent synthetic utterance to the authentication system in the

voice of the target speaker. In practice, however, current state-of-the-art text-to-speech synthesis is not quite capable of producing such natural sounding utterances. In other words, synthetic speech produced by current TTS systems still sounds far from natural and is easily distinguished from genuine human speech by the human ear. Does this mean, however, that TTS speech could not deceive an authentication system based on automatic speaker recognition? To answer this question, it needs to be examined how different speaker recognition systems actually work.

As shown in Table 1, there are three types of speaker recognition systems that are distinct by the types of speech patterns each examines in order to determine the similarity of the unknown speech and the target speech. The most common type of speaker recognition system looks at speaker differences at the individual sound level. A second type of speaker recognition system examines the sequences of speech sounds, which form words, and a third type also analyzes higher-level information such as intonation, choice of words, choice of sentence structure or even semantic or pragmatic content of the utterances in question [4].

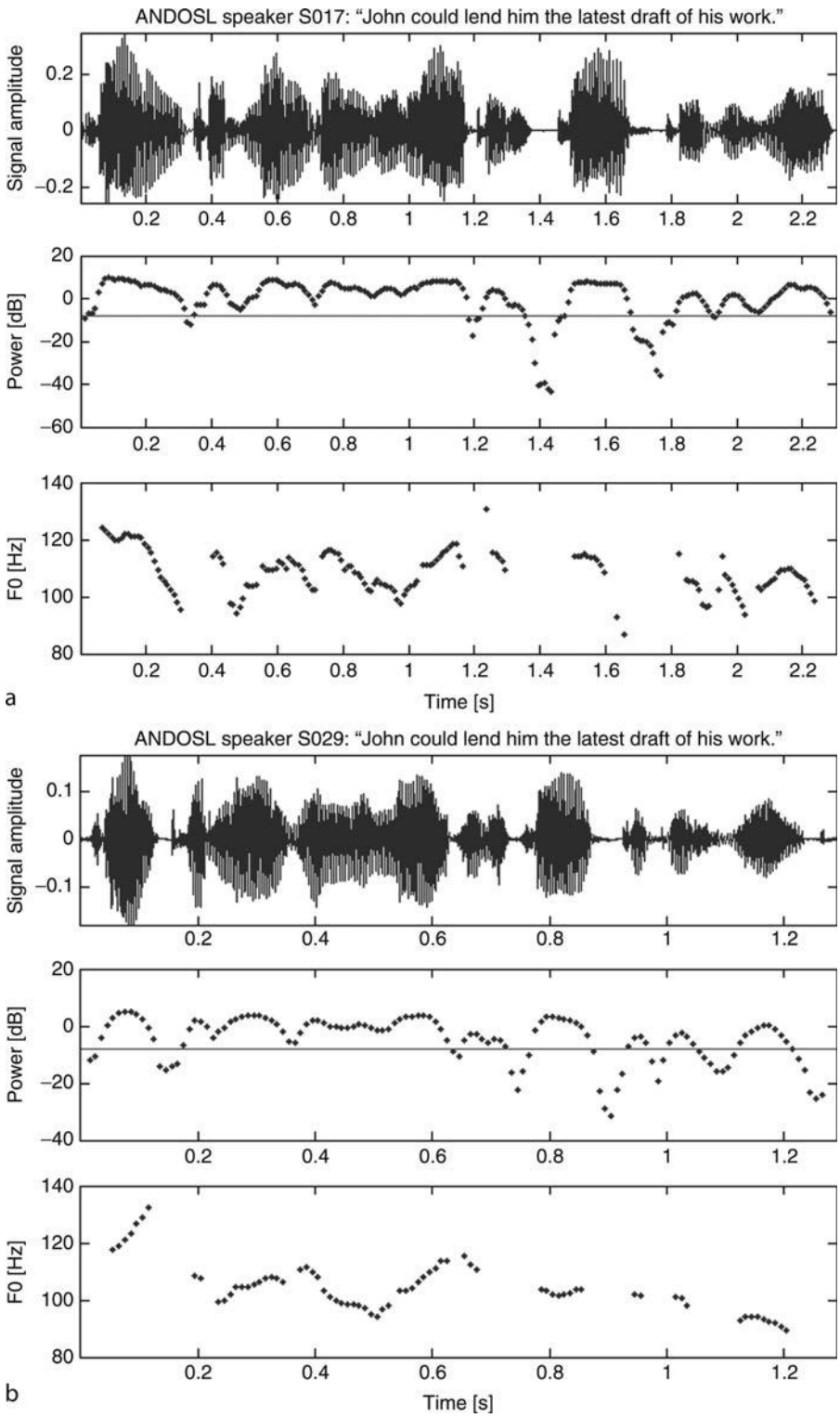
Speech processing invariably segments a speech signal into small chunks, or "frames", which correspond in duration roughly to short speech sounds, say about 10–30 ms. For each frame, features are extracted from the speech signal, such as a spectrum or a cepstrum or a mel-frequency cepstrum (MFC) [5]. These extracted features serve as the basis for the comparison between the unknown speech and the target speech. The first type of speaker recognition system independently compares the features of each frame of the unknown speech signal with the model of the target speaker. This is done independently for each frame and without considering the speech sounds immediately preceding or succeeding the given frame.

The second type of speaker recognition system takes into account the likelihood of *sequences* of speech sounds, rather than individual speech sounds, when comparing the unknown speech signal with the model of the target speaker. For example, the sound sequence /the/ would be more likely for a speaker of English than the sound sequence /eth/. The third type of system takes into account higher-level features, i.e., the variation of features over time such as the intonation pattern of a sentence, as it manifests itself through the functions of loudness and pitch over time. Such authentication systems typically operate on much longer passages of speech, for example, to segment a two-way telephone conversation or the proceedings in a court of law into the turns belonging to the different speakers. Figure 4 shows an example of two speakers pronouncing the same sentence with quite different intonation.

It is easy to see that a context-free authentication system is prone to be attacked successfully by a very simple synthesizer, namely one that produces a few seconds of only a single speech sound. For example, an attacker could reproduce a single frame of, say, the sound "a" of the target speaker and play this frame

Liveness Assurance in Voice Authentication. **Table 1** Types of speaker authentication methods

Type of speaker recognition system	Training/Enrolment	Testing	Typical method
Recognizes individual speech sounds (context-free)	A set of speech sounds typical for the target speaker is collected and becomes the "model" for the target speaker	Each speech sound is individually compared with the target speaker "model"	Gaussian MixtureModel (GMM)
Recognizes sequences of speech sounds (context-sensitive)	In addition to the individual sounds, the speaker model represents the sequences of speech sounds that are typical for the target speaker	The entire utterance is compared with the target speaker model for both individual sounds and sound sequences	Hidden Markov Model (HMM)
Recognizes higher-level features (intonation, word choice, syntax etc.)	In addition to sound sequences, the speaker model represents words, sentence structures and intonation patterns typical for the target speaker	Similarity of sounds and sound sequences is combined with similarity of word sequences and intonation patterns	Information fusion of GMM and/or HMM with higher-level information sources



Liveness Assurance in Voice Authentication. Figure 4 Two male speakers from the Australian National Database of Spoken Language (ANDOSL), speaking the same sentence with distinctly different intonation: audio signal, and power and fundamental frequency (F0) contours. Speaker S017 produced the word *John* with falling F0, while Speaker S029 produced the same with rising F0.

repeatedly for a second or two in order to "convince" an authentication system of this type that the "aaaaaaa. . ." sound represents the natural voice of the target speaker. This is because each frame is assessed independently as being similar to the "a" sound of the target speaker, irrespective of the fact that the sequence of "a" sounds does not represent a likely speech pattern of the target voice.

A context-sensitive authentication system, on the other hand, requires a speech synthesizer to reproduce entire sound sequences that are sufficiently similar to sound sequences produced by the target speaker. This means that the individual sounds, produced by the synthesizer, must be similar to sounds of the target speaker and the sound sequences must be structured in a similar way to those of the target speaker. This is a proposition that is far more difficult, although not impossible, to achieve with current-state-of-the-art speech synthesizers. Furthermore, if the speaker authentication system also considers the intonation pattern and higher-level features such as choice of words and grammatical constructs, an attacker who tries to impersonate a target speaker using a TTS synthesizer, would require a system that is beyond the capabilities of the technology at the time of writing.

Multimodal Liveness Assurance

The assurance that a voice biometric is delivered live at the time and place of authentication can be enhanced considerably by complementing the voice modality with a second modality. In the simplest case, this could be the visual modality provided by a human observer who can assure that the voice biometric is actually provided by the person seeking authentication and that person is not using any device to play back a recorded or synthesized voice sample.

In an automatic voice authentication system, similar assurance of liveness can be achieved by combining the voice modality with a face recognition system. Such a system has a number of advantages. Firstly, the bimodal face-voice approach to authentication provides two largely independent feature sets, which, when combined appropriately, can be expected to yield better authentication than either of the two modalities by itself. Secondly, the bimodal approach will add robustness to the system when either modality is affected by difficult environmental conditions. In the case of bimodal face-voice authentication, it is particularly useful to fall back on the complementary face recognition facility when the voice recognition modality breaks down due to high levels of surrounding noise, competing speakers or channel variability such as that caused by weak cell phone reception. In such situations, the face recognition modality will be able to take over and hence provide enhanced robustness for the combined system.

A similar consideration applies, of course, when the combined face-voice authentication system is viewed from the perspective of the face recognition modality, which may equally break down in difficult environmental conditions such as adverse lighting. In this case, too, the overall robustness of the authentication system is preserved by the combination of the two modalities voice and face, each of which is affected differently and largely independently by environmental factors.

However, the most important advantage of a bimodal face-voice authentication system for the assurance of liveness is the fact that the articulator movements, mainly of the lips, but also of the tip of the tongue, jaw, and cheeks are mostly observable and correspond closely to the particular speech sounds produced. Therefore, it is possible when observing a bimodal audio-video signal of the speaking face to ascertain whether the facial dynamics and the sequence of speech sounds are mutually compatible and synchronous. To a human observer it is quite disconcerting when this is not the case, for example, with an out-of-sync television signal or with a static facial image when the speaker is heard saying something, but the lips are not seen to be moving. In the field of audiovisual speech recognition, the term "viseme" has been coined as the visual counterpart of the "▶ phoneme", which denotes a single speech sound. The visemes /m/, /u/, and /d/ (as in the word "mood"), for example, first show the speaker's lips spread and closed (for /m/), then protruded and rounded (for /u/), and finally spread and slightly open (for /d/). It is therefore possible to detect whether the corresponding sequences of visemes and phonemes of an utterance are observed in a bimodal audio-video signal and whether the observed viseme and phoneme sequences are synchronous.

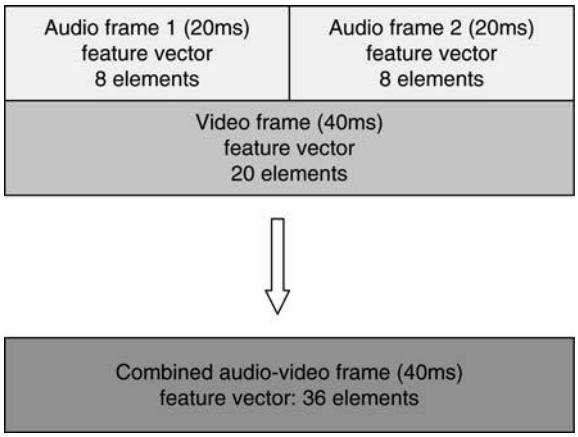
In order for the synchrony of the audio and video streams to be ascertained, the two modalities must be combined appropriately. Multimodal authentication

systems employ different paradigms to combine, or "fuse", information from the different modalities. Modality fusion can happen at different stages of the authentication process. Fusing the features of the different channels immediately after the feature extraction phase is known as "feature fusion" or "early fusion". In this paradigm, all comparisons between the unknown sample and the client model as well as the decision making are based on the combined feature vectors. The other possibility is to fuse information from the two modalities after independent comparisons have been made for each modality. Such paradigms are known as ▶ score fusion, ▶ decision fusion or ▶ late fusion.

For liveness assurance by means of bimodal face-voice authentication, it is necessary to apply an early fusion stratagem, i.e., to fuse the two modalities at the feature level [6]. If the two modalities were fused late, i.e., at the score or decision level, analysis of the video of the speaking face would yield one decision on the speaker's identity and analysis of the audio of the utterance would yield another decision on the speaker's identity. The two processes would run independently of each other with no connection between them that would allow the checking for the correspondence and synchrony of visemes and phonemes [7].

Therefore, the features that are extracted from the audio signal on a frame-by-frame basis – usually at an audio frame rate of about 40–100 frames per second – must be combined with the features that are extracted from the video signal – usually at the video frame rate of 25 or 30 frames per second. An example of how the differing frame rates for the audio and video signals can be accommodated is shown in Fig. 5, where the audio frame rate is 50 frames per second, the video frame rate is 25 frames per second, and the combined audiovisual feature vector comprises the audio feature vectors of two consecutive audio frames, combined with the single video vector of the synchronous video frame.

The combined audiovisual feature vectors will then reveal whether the audio and video streams are synchronous, for example, when the combined audiovisual feature vectors contain the sequence of visemes /m/, /u/, and /d/ and likewise the sequence of phonemes /m/, /u/, and /d/. In contrast, if one of the combined audiovisual feature vectors were to contain the visual information for the viseme /m/ and at the same time the audio information for the phoneme /u/, the combined feature vector would indicate that the audio and video streams do not represent a corresponding synchronous representation of any speech sound.



Liveness Assurance in Voice Authentication. Figure 5 Feature fusion of two consecutive 20ms audio feature vectors with the corresponding 40ms video feature vector. Before fusion, the audio vectors have been reduced to 8 dimensions each, and the video vector has been reduced to 20 dimensions. The combined feature vector has 36 dimensions.

The proper sequencing of visemes and phonemes is usually ascertained by representing the audio-visual speech by Hidden Markov Models (HMM), which establish the likelihoods of the different combined audiovisual vectors and their sequences over time [8]. It is therefore possible to ascertain whether the audio and video components of a combined audio-video stream represent a likely live utterance. Therefore, an attacker who attempts to impersonate a target speaker by means of a recorded speech utterance and a still photograph of the target speaker will be thwarted because the system will recognize the failure of the face to form the corresponding visemes that should be observed synchronously with the phonemes of the utterance. Similarly, such a system will thwart an attack by an audiovisual speech synthesis system, unless the synthesizer can generate the synthetic face and the synthetic voice in nearly perfect synchrony.

Related Entries

▶ Biometric Vulnerabilities: Overview
▶ Security and Liveness, Overview
▶ Voice Authentication

References

1. Furui, S.: Cepstral analysis techniques for automatic speaker verification. IEEE Trans. Acoust., Speech and Signal Processing ASSP-**29**, 254–272 (1981)
2. Bimbot, F., Bonastre, J.-F., Fredouille, C., Gravier, G., Magrin-Chagnolleau, I., Meignier, S., Merlin, T., Ortega-García, J., Petrovska-Delacrétaz, D., Reynolds, D.A.: A tutorial on text-independent speaker verification, EURASIP J. Appl. Signal Processing **2004**(4), 430–451 (2004)
3. Matsui, T., Furui, S.: Speaker adaptation of tied-mixture-based phoneme models for text-prompted speaker recognition. In: Proceedings of International Conference on Acoustics, Speech and Signal Processing, Adelaide, pp. I-125–128. IEEE Publisher, New York (1994)
4. Reynolds, D., Andrews, W., Campbell, J., Navratil, J., Peskin, B., Adami, A., Jin, Q., Klusacek, D., Abramson, J., Mihaescu, R., Godfrey, J., Jones, D., Xiang, B.: The SuperSID Project: Exploiting high-level information for high-accuracy speaker recognition. In: Proceedings of International Conference on Acoustics, Speech and Signal Processing, Hong Kong, pp. IV-784–787. IEEE Publisher, New York (2003)
5. Huang, X., Acero, A., Hon, H.-W.: Spoken language processing, Prentice Hall, Upper Saddle River, NJ (2001)
6. Chetty, G., Wagner, M.: Investigating feature-level fusion for checking liveness in face-voice authentication. In: Proceedings of eighth IEEE Symposium on Signal Processing and its Applications, Sydney, pp. 66–69. IEEE Publisher, New York (2005)
7. Bredin, H., Chollet, G.: Audiovisual speech synchrony measure: Application to biometrics. EURASIP J. Adv. Signal Process. **2007** (1), 1–11 (2007)
8. Chetty, G., Wagner, M.: Speaking faces for face-voice speaker identity verification. In: Proceedings of Interspeech-2006 – International Conference on Spoken Language Processing, Paper Mon3A1O-6, Pittsburgh. International Speech Communication Association (2006)

Liveness Detection

Liveness detection is a functionality that determines whether the presented biometric sample (e.g., finger, hand, or iris) is originated from a live body. This functionality is considered to be one of the key security measures that improve the reliability of a biometric system because it enables the system to reject artifacts to be enrolled and ensure that no forged sample is accepted.

► Finger Vein Reader

Liveness Detection

In biometric systems, the goal of liveness testing is to determine if the biometric being captured is an actual measurement from the authorized, live person who is present at the time of capture. While biometric systems may have an excellent performance and improve security, previous studies have shown it is not difficult to spoof biometric devices through fake fingers, high resolution images or video, contact lenses, etc. Even though biometric devices use physiologic information for identification/verification purposes, these measurements rarely indicate liveness. Liveness detection reduces the risk of spoofing by requiring a liveness signature in addition to matched biometric information. Methods can include medical measurements such as pulse oximetry, electrocardiogram, or odor. In a few cases, liveness information is inherent to the biometric itself, i.e., the biometric cannot be captured unless the user is live, e.g., electrocardiogram as a biometric. While liveness algorithm makes spoofing more difficult, they need to be considered as components of a biometric system which bring with it performance characteristics, as well as factors such as ease of use, collectability, user acceptance, universality, spoof-ability, permanence, and, in some cases, even uniqueness. No system is perfect in its ability to prevent spoof-attacks. However, liveness algorithms can reduce this vulnerability to minimize the risk of spoofing.

► Anti-spoofing
► Liveness Detection: Fingerprints
► Liveness Detection: Iris
► Liveness Detection: Fingerprint

Liveness Detection: Fingerprint

Stephanie A. C. Schuckers
Clarkson University, Potsdam, New York, USA

Synonyms

Anti-spoofing; Vitality

Definition

In biometric systems, the goal of liveness testing is to determine if the biometric being captured is an actual measurement from the authorized, live person who is present at the time of capture. While fingerprint systems may have an excellent performance and improve security. Previous studies have shown it is not difficult to make molds of latent fingerprints left by legitimate users and to create fake fingers made from Play-Doh, gelatin, and silicone materials to fool a variety of fingerprint scanners, termed spoofing. Liveness detection reduces the risk of spoofing by requiring a liveness signature, in addition to matched biometric information. Methods can be divided into hardware and software categories. Hardware methods include measurements like pulse oximetry, electrocardiogram, or odor, while software based measurements use additional processing of the biometric information itself to isolate liveness signatures like perspiration and deformation. While liveness algorithm makes spoofing more difficult, they need to be considered as components of a biometric system, which bring with it performance characteristics along with factors such as ease of use, collectability, universality, spoof-ability, permanence, and in some cases, even uniqueness. No system is perfect in its ability to prevent spoof-attacks. However, liveness algorithms can reduce this vulnerability to minimize the risk of spoofing.

Fingerprints are graphical ridge-valley patterns from human fingers. Fingerprint recognition is a widely used and efficient technique for biometric authentication. While fingerprint systems may have excellent performance and improve security, previous studies have shown it is not difficult to make molds of latent fingerprints left by legitimate users and to create fake fingers made from Play-Doh, gelatin and silicone materials to fool a variety of fingerprint scanners [1, 2]. The most famous of which is the work by Matsumoto and colleagues. In the reports, two different techniques were used to create a mold. The first technique directly used a subject's finger to create the mold in free molding plastic, whereas the second technique involved making a mold from a latent fingerprint image. Casts were made of gelatin material and termed 'gummy fingers'. Verification rates of gummy fingers ranged from 68 to 100%. For method of creating a cast from residual fingerprints, all fingerprint systems were able to enroll the spoof finger and verify more than 67% of the attempts. Similar results have been obtained on subsequent studies with various materials including silicon, clay, and Play-Doh [1, 2], and one study which looked at cadaver fingers [2]. Currently, International Biometric Group with sponsorship from Financial Services Technology Consortium (FSTC) is hosting an effort to conduct spoof trials with vendor volunteers called SPOOF 2007.

It should be noted that vulnerability to spoofing is not assessed as part of the false accept ratio, a typical assessment measure of biometric devices. A false accept is when a submitted sample is incorrectly matched to a template enrolled by another user. This only refers to a zero-effort attempt, i.e., an unauthorized user making an attempt with their own biometric to gain access to a system. If the false accept ratio is kept low, then the probability of specific user *with criminal intent* matching another template is very low. The false accept ratio does not give information on the vulnerability of a system to spoof attacks.

Even though biometric devices use physiologic information for identification/verification purposes, these measurements rarely indicate liveness. The goal of liveness testing is to determine if the biometric being captured is an actual measurement from the authorized, live person who is present at the time of capture. Overview of liveness approaches are described in [2–5]. Performance of fingerprint liveness to separate live and spoof fingers is measured by live false reject rate and spoof false accept rate. Equal error rate between these two measures and receiver operating characteristic curves can also be used as described in Biometric Security Overview. Marcialis et al., provides a table which compares datasets used for testing and performance of liveness approaches.

Methods to measure liveness fall into several categories. In 2007 [6], a taxonomy is presented whereby methods are divided into software and hardware-based. A similar division is suggested, but also consider an additional category where liveness is inherent to the biometric, i.e., it must be present in order to capture the biometric [2]. In the first liveness is captured through additional hardware integrated with the fingerprint sensor. For first category software-based techniques, involves further processing of the biometric signature to obtain liveness information. For example, this may mean extracting perspiration information from fingerprint image. The second software based approach is where liveness is an inherent part of the biometric, in other words, the biometric cannot be

captured unless the subject is alive. An example for this category is the electrocardiogram, which has been suggested as a biometric [7] and where liveness is inherent to collection of this biometric. Liveness in most cases is not inherent to be able to measure a fingerprint biometric. Most systems that consider liveness in fingerprint do so through additional software or hardware processing. Electrocardiogram might be considered a special case as it has been suggested as an additional measurement to fingerprint recognition so it can be considered as hardware liveness approach and it may be potentially inherent to the biometric if the electrocardiogram is used as a biometric.

Hardware

The first method uses extra hardware to acquire life signs. Previously developed approaches measure fingertip temperature, pulse, pulse oximetry, blood pressure, electric resistance, odor, multi-spectral information, or electrocardiogram (e.g., [8, 7, 9, 10]). These methods require dedicated hardware integrated with the fingerprint system. Electrocardiogram is the electrical measurement of the heart collected through electrodes on two skin contact points on the body which need to be on opposite sides of the heart (e.g., two hands, hand and foot). Pulse oximetry is the measurement of the oxygen content of the blood through the comparison of the absorption of two wavelengths of light by the blood. This measurement requires a LED and photodetector on opposite sides of the finger and typically needs to be shielded from ambient light. This absorption also varies as the heart beats and can be a measure of pulse, and therefore may require a few seconds to compute to record one or two complete heart beat cycles. A critical component to hardware-based approaches is how the additional hardware is integrated with the fingerprint sensor. It should be integrated in such a way that it cannot be spoofed with any live finger in combination with a spoof.

The following paragraph describes two fingerprint sensors, multispectral and ultrasound, which naturally capture liveness information. They are placed here in the hardware category, because these approaches, while commercially viable, require purchase of a specific scanner and are not applicable to standard fingerprint readers. One commercially available fingerprint sensor (Lumidigm, USA) uses a multispectral sensor, from which multiple wavelengths of light and different polarizations allow new data to be captured, which is unavailable from a conventional optical fingerprint reader. Based on the multiple spectral images, they have developed a spoof detection method [10]. Similarly, ultrasound measurements have been suggested as a way to measure fingerprint images (Optel, Poland). While fingerprint measured by ultrasound might be able to image a spoof or cadaver fingerprint itself, using additional information from the ultrasound measurement would likely be capable of separating live from spoof images. Both approaches most likely need additional processing from the fingerprint image itself to determine liveness.

Software

The second method uses the information already present in the fingerprint image to detect life signs, for example, skin deformation, pores, power spectrum, or perspiration pattern.

Skin deformation and elasticity. Skin deformation technique uses the information regarding how the fingertip's skin deforms when pressed against the scanner surface [11–14]. The studies show that when a real finger moves on a scanner surface, it produces a significant amount of non linear distortion. However, fake fingers are more rigid than skin and the deformation is lower even if they are made of highly elastic materials. One approach quantifies this considering multiple frames of clockwise motion of the finger [12]. The performance of this method has an equal error rate of 11.24% using 45 live subjects and 40 fake fingers. A study by Zhang et al. [14] uses a thin-plate spline distortion model over multiple frames, while the finger is moved and resulted 4.5% EER in a dataset of 120 fake fingerprints from silicon from 20 individuals. Another method considers the deformation in a single image compared to a template [11]. This study achieved 82% for a small dataset.

Perspiration pattern. Previously, our laboratory has demonstrated that perspiration can be used as a measure of liveness detection for fingerprint biometric systems. Unlike spoof and cadaver fingers, live fingers demonstrate a distinctive spatial moisture pattern when in physical contact with the capturing surface

of the fingerprint scanner. The pattern in fingerprint images begins as 'patchy' areas of moisture around the pores spreading across the ridges over time. Image/signal processing and pattern recognition algorithms have been developed to quantify this phenomenon using wavelet and statistical approaches [15–17]. These approaches require two time-series images, which might not be convenient for the users. Other methods to quantify this phenomenon have been developed for a single image [18]. Performance has achieved approximately 10% live/spoof EER for earlier papers on a dataset of 80 spoof, 25 cadaver, and 58 live images to perfect separation in later papers on this small dataset [16].

Characteristics of spoof and live images. A natural extension to the specific categories above is to begin to assess the characteristics that define live and spoof fingers, which cover a broad range [5, 13, 19–21]. These include image power spectrum that reveals stamp fabrication process [5], noise residue in the valleys due to spoof material [19, 21], and combinations of multiple factors, for example, fusion of perspiration and deformation features [13].

Image power spectrum has been considered as an effective feature for vitality detection [5, 20]. The difference between live and spoof images is mainly due to the stamp fabrication process, which causes an alteration of frequency details between ridge and valleys. The Fourier transform feature can quantify the difference in terms of high frequency information loss for fake fingers. This approach is tested for a single scanner and silicone spoof material with average spoof/live EER of 2.4% on a dataset of 720 fake and 720 live images from 36 individuals [6] and for gelatin and silicon with an average of 23% EER for a dataset of 900 fake and 450 live images from 30 individuals [20].

In other study [13], a sequence of images is used to measure skin elasticity, but some of the measures may be capturing perspiration information as described above. No special motion is required for the finger. They achieve results of 4.78% on a dataset of 470 spoof images from 47 spoof casts and 300 live images from 15 individuals. In a second study, fusion of multiple features, two based on perspiration signal and two based on skin elasticity, was performed in 2007 [22]. Result showed 4.49% EER on the same dataset.

Liveness Algorithm Framework

Fingerprint liveness algorithms can fall into types described above (hardware, software, and inherent). Other factors that separate liveness algorithms include (1) dynamic/static, (2) user training, and (3) binary/user specific. Table 1 compares five fingerprint liveness algorithms within the context of this framework.

- *Dynamic or static*: Liveness algorithms may require only one frame or rely on multiple frames to measure the dynamic nature of the system to detect liveness [5]. For example, many of the perspiration proposed approaches require more than one image [15], although recent work has used one image [18]. Other dynamic approaches are related to deformation [12–14]. Note that pulse oximetry do not require multiple fingerprint image frames, however, they may require more time to record one or more full heart cycles.
- *User training*: Some liveness algorithms rely on specific user actions to determine liveness. This may include a procedure (deformation changes due to rotating the finger), which require user training [12, 14].
- *Binary (live/spoof) versus user specific*: Liveness algorithms can be made general across all subjects, that is, the same algorithm is used for all subjects to determine liveness producing a binary result: live or non live (Fig. 1). Other approaches can be made subject specific, that is, a liveness algorithm is

Liveness Detection: Fingerprint. Table 1 Liveness algorithm types and factors[a]

	Hardware/ Software	Multiple/ Single	Binary/ User specific	User training
Perspiration	S	M/Si	B/US	None
Pulse oximetry	H	–	B	None
Multi-spectral	H	Si	B/US	None
Deformation	S	M/Si	B	UT or none
ECG	H	–	B/US	UT

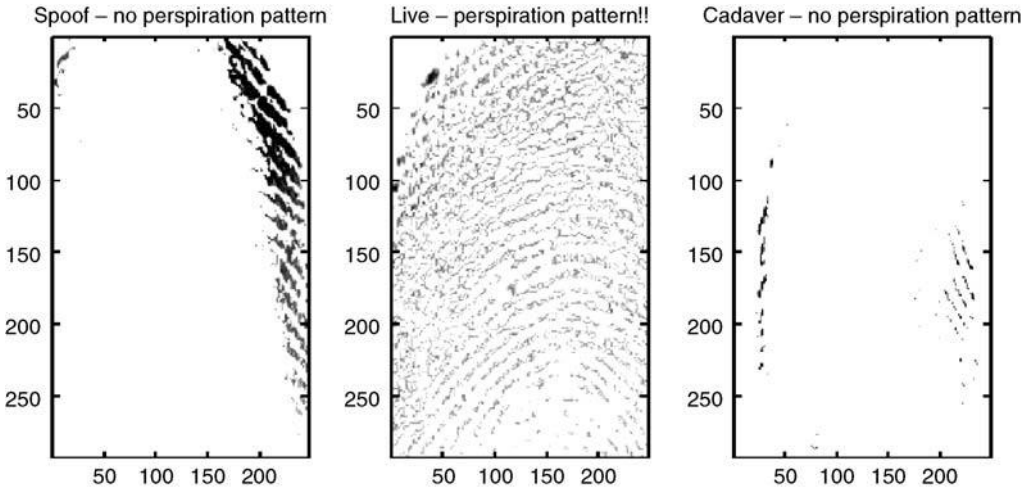
[a]*H* Hardware; *S* Software; *M* Multiple; *Si* Single; *B* Binary; *US* User Specific; *UT* User Training; – indicates not applicable

imbedded as part of the biometric template. For example, work has been shown for storing a perspiration pattern specific to an individual [23]. While not specifically mentioned for the multi-spectral fingerprint scanner (Lumidigm, USA), it is possible that a medical spectroscopy-based liveness approach could be user specific. Electrocardiogram can also be user-specific, that is, used as a biometric [7] (Fig. 2).

Liveness Detection: Fingerprint. **Figure 1** Example of live and non-live fingerprints captured by Capacitive DC scanner: (**a**) live finger; (**b**) spoof finger made from Play-Doh; (**c**) spoof finger made from gelatin; (**d**) cadaver finger.

Other characteristics for evaluating biometrics systems, such as ease of use, collectability, user acceptance, universality, uniqueness, permanence, and spoof-ability, need to be considered before implementing a liveness algorithm. These were described in the Biometric Security Overview chapter. Table 2 considers the same liveness algorithms from Table 1 within the context of this framework.

- *Ease of use*: Some liveness approaches may be easier to use. For example, fingerprint deformation approach that requires a specific rotation procedure may be considered more difficult to use [12, 13]. Lumidigm approach for spectroscopy where liveness is collected as part of the biometric collection itself may be considered easier to use.
- *Collectability*: The hardware, equipment setup, and relationship to the user impacts the collectability of the liveness algorithm. For example, approaches that may be more difficult to collect include the electrocardiogram, which requires two points of contact on opposite sides of the body or pulse oximetry, where the finger must be enclosed to protect from ambient light. In comparison, approaches that use the traditional biometric



Liveness Detection: Fingerprint. **Figure 2** Perspiration patterns. Spoof, live, and cadaver patterns are shown from left to right. The perspiration pattern is the reconstruction of the isolated wavelet coefficients obtained from two fingerprint images in time, by the algorithm described [4].

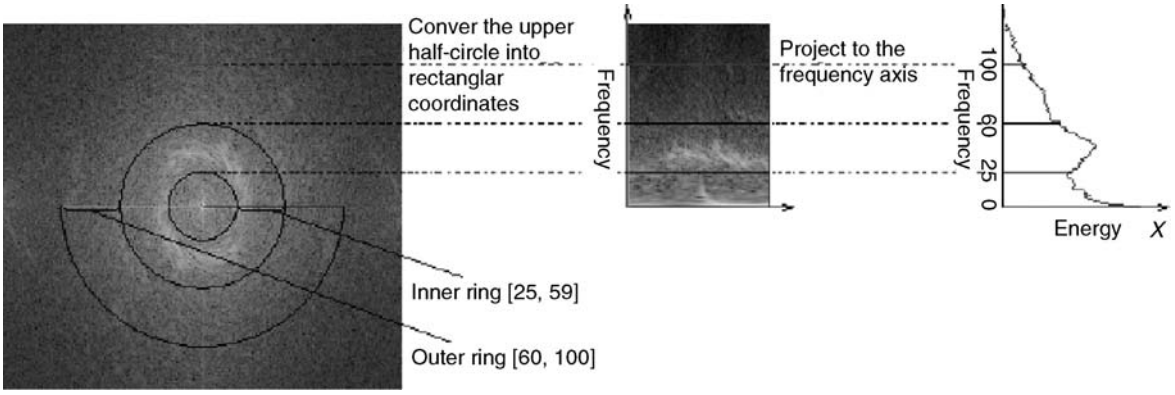
equipment for measurement of liveness might be considered easier to collect.

- *User acceptance*: For fingerprint liveness, approaches with low user acceptance are ones that are more likely to be linked with medical conditions due to privacy concerns (electrocardiogram, pulse oximetry, and multi-spectral) (Fig. 3).
- *Universality*: Obviously all authorized users should be live when presenting their biometric; however, the liveness signature may be difficult to measure in some subjects. For example, perspiration in fingerprint images may be difficult to measure in individuals with very dry skin, which is also a problem with measuring the fingerprint image itself.
- *Uniqueness*: For liveness approaches, which are inherent to the biometric, this factor is critical. However, as mentioned above, electrocardiogram in combination with fingerprint would not need uniqueness as a characteristic, whereas, the electrocardiogram alone may need further research to address uniqueness [7].
- *Permanence*: Permanence typically refers to permanence of the specific biometric pattern over time. Similar to above, this more directly applies to liveness approaches, which are inherent to the biometric, where the biometric/liveness signature may vary over time. For example, in the initial work introducing perspiration patterns as a unique liveness pattern, only 3 months were considered [23]. It is unknown if these patterns persist beyond that period. Electrocardiogram may also have difficulties with permanence as the electrocardiogram is impacted by health conditions [7].
- *Spoof-ability*: Spoof-ability considers the possibility that the liveness mechanism which is put in place to protect the system from spoofing can be spoofed. For example, in the case of pulse oximetry, it may be possible to spoof with a clear spoof, which allows transmission of the light needed to make

Liveness Detection: Fingerprint. Table 2 Liveness algorithm characteristics[a]

	Ease of use	Collectability	User acceptance	Universality	Uniqueness	Permanence	Spoof-ability
Perspiration	H	H	H	M	L	M	M
Pulse oximetry	L	L	L	H	–	–	H
Multi-spectral	H	H	M	H	–	–	L
Deformation	L	L	H	M	–	–	M
ECG	L	L	L	H	L	H	H

[a]*H* High; *M* Medium; *L* Low; –indicates not applicable



Liveness Detection: Fingerprint. Figure 3 Spectral image of the fingerprint, the ring pattern, and the band-selected frequency analysis from [20].

the pulse oximetry measurement. This goes beyond the performance of the liveness algorithm described above, because it requires assessment of spoofing approaches that have yet to be replicated in the database used to test the liveness algorithm.

Summary

In summary, liveness systems are being suggested to reduce the vulnerability due to spoofing. Liveness measures have an inherent performance, that is, ability to separate spoof and live attempts. In addition, liveness algorithms have other factors and considerations including ease of use, collectability, user acceptance, universality, uniqueness, permanence, and spoof-ability. One factor, which is difficult to measure is spoof-ability, the possibility that the liveness measure can be spoofed. In this chapter, the term liveness is used, fully acknowledging that it is not a perfect system and that it is not possible to recreate all possible spoof attempts for a system. Furthermore, there may be measurements, which rule out specific spoofs but cannot be shown to absolutely measure liveness. For example, algorithms may be designed which may readily detect silicon, but not gelatin, spoof images. In summary, it is unlikely that any system will perfectly measure liveness and be spoof-proof. Liveness may be boiled down to an attempt to stay one step ahead of those intending to defeat the system through spoof attacks. Methods such as liveness or antispoofing are critical to the security and credibility of biometric systems to protect them from security vulnerabilities to the degree needed for a particular application.

Related Entries

► Liveness Iris
► Security and Liveness

References

1. Matsumoto, T., Matsumoto, H., Yamada, K., Hoshino, S.: Impact of artificial 'gummy' fingers on fingerprint systems. Proc. SPIE **4677**, 275–289 (2002)
2. Schuckers, S.A.C.: Spoofing and anti-spoofing measures. Inform Security Tech Rep **7**(4), 56–62 (2002)
3. Valencia, V., Horn, C.: Biometric liveness testing. In: Woodward J.D. Jr., Orlans, N., Higgins M.R.T. (eds.) Biometrics. McGraw-Hill, Osborne Media, New York (2002)
4. Schuckers, S.A.C., Abhyankar, A.: A wavelet based approach to detecting liveness in fingerprint scanners. Paper presented at the Proceedings of Biometric Authentication Workshop, ECCV, Prague, May 2004
5. Coli, P., Marcialis, G.L., Roli, F.: Power spectrum-based fingerprint vitality detection. Paper presented at IEEE Workshop on Automatic Identification Advanced Technologies AutoID pp. 169–173 (2007)
6. Coli, P., Marcialis, G.L., Roli, F.: Vitality detection from fingerprint images: A critical survey. Adv. Biometrics **4642**, 722–731 (2007)
7. Biel, L., Pettersson, O., Philipson, L., Wide, P.: ECG analysis: A new approach in human identification. IEEE Trans. Instrum. Meas. **50**(3), 808–812 (2001)
8. Kallo, P., Kiss, I., Podmaniczky, A., Talosi, J.: Detector for recognizing the living character of a finger in a fingerprint recognizing apparatus. Dermo Corporation, U.S. Patent No. 6,175,64. 16 Jan 2001
9. Baldisserra, D., Franco, A., Maio, D., and Maltoni, D.: Fake fingerprint detection by odor analysis. Paper presented at the proceedings of International Conference on Biometric Authentication (ICBA06), Hong Kong, Jan 2006
10. Nixon, K.A., Rowe, R.K.: Spoof detection using multispectral fingerprint imaging without enrollment. Paper presented at the Proceedings of Biometrics Symposium (BSYM2005), Arlington, VA, 19–21 Sept 2005
11. Chen, Y., Jain, A., Dass, S.: Fingerprint deformation for spoof detection. Paper presented at the Proceedings of Biometrics Symposium (BSYM2005), Arlington, VA, 19–21 Sept 2005
12. Antonelli, A., Cappelli, R., Maio, D., Maltoni, D.: Fake finger detection by skin distortion analysis. IEEE Trans. Inform. Forensics Security **1**(3), 360–373 (2006)
13. Jia, J., Lianhong, C., Kaifu, Z., Dawei, C.: A new approach to fake finger detection based on skin elasticity analysis. Adv. Biometrics **4642**, 309–318 (2007)
14. Zhang, Y., Tian, J., Chen, X., Yang, X., Shi, P.: Fake finger detection based on thin-plate spline distortion model. Adv. Biometrics **4642**, 742–749 (2007)
15. Derakhshani, R., Schuckers, S., Hornak, L., O'Gorman, L.: Determination of vitality from a non-invasive biomedical measurement for use in fingerprint scanners. Pattern Recogn. **17**(2), 383–396 (2003)
16. Schuckers, S.A.C, Derakhshani, R., Parthasardhi, S., Hornak, L.A.: Liveness detection in biometric devices. In: Electrical Engineering Handbook, Chapter 26, 3rd edn. CRC, Boca Raton (2006)
17. Parthasaradhi, S., Derakhshani, R., Hornak, L., Schuckers, S.A.C.: Time-series detection of perspiration as a liveness test in fingerprint devices. IEEE Trans. Systems Man Cybern C Appl Rev **35**, 335–343 (2005)
18. Tan, B., Schuckers, S.: Liveness detection using an intensity based approach in fingerprint scanner. Paper presented at Proceedings of Biometrics Symposium (BSYM2005), Arlington, VA, 19–21 Sept. 2005

19. Moon, Y.S., Chen, J.S., Chan, K.C., So, K., Woo, K.C.: Wavelet based fingerprint liveness detection. Electron. Lett. **41**(20), 1112–1113 (2005)
20. Jin, C., Kim, H., Elliott, S., Liveness detection of fingerprint based on band-selective fourier spectrum. Information Security Cryptol –ICISC, vol. 4817, pp. 168–179. Springer, Berlin (2007)
21. Tan, B., Schuckers, S.: A new approach for liveness detection in fingerprint scanners based on valley noise analysis. J. Electron. Imag. **17**(1), 011009-1–011009-9 (2008)
22. Jia, J., Lianhong, C.: Fake finger detection based on time-series fingerprint image analysis. In: Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues, vol. 4681, pp. 1140–1150. Springer, Berlin (2007)
23. Abhyankar, A., Schuckers, S.A.C.: Characterization, similarity score, and uniqueness of fingerprint perspiration patterns. In: Kanade et al. (eds.) et al.: Proceedings of Audio- and Video-Based Biometric Person Authentication: 5th International Conference. Lecture Notes in Computer Science, vol. 3546, pp. 860–868. Springer, Berlin (2005)

Liveness Detection: Iris

Bori Toth
Deloitte & Touche LLP, London, UK

Synonyms

Anti-spoofing; Spoofing countermeasures; Spoof-resistance; Vitality tests

Definition

Iris liveness detection techniques are designed to counteract physical ▶ spoofing attacks launched against iris recognition systems. Such attacks include the use of photographs, video recordings, printed contact lenses etc. Iris liveness detection mechanisms aim to ascertain that iris images acquired were acquired from a live and authorized user present at the time of transaction.

Introduction

Iris recognition systems are among the most accurate biometric systems available today. Additionally, the iris is an internal organ which makes it more robust to spoofing attacks when compared to some of the other biometric technologies, especially fingerprinting. This is mainly due to two reasons. First, unlike DNA and fingerprints, individuals do not leave traces of their irises behind which could be lifted and copied. Second, it is more difficult to manipulate an internal organ than to disguise an external body part such as the face. Nevertheless, all technologies have inherent weaknesses which can be exploited, including iris recognition.

On the one hand, privacy experts argue that biometric information is private. Additionally, the growing uptake of large-scale biometric systems worldwide intensifies the fear of hackers stealing biometric information from centralized databases. On the other hand, it is a fact that the acquisition of biometric information is much easier than breaking databases. Facial images are being taken via public and private CCTV systems day and night: think of ATMs, petrol stations, banks, and airports just to name a few places were camera surveillance systems are usually in 24/7 operation. The situation is similar for most other biometric traits: we leave our fingerprints and DNA behind on every surface we touch and phone-based service providers usually record our voices during phone calls [1]. In the case of iris recognition, a high resolution image of someone's eye can be sufficient to make the technology work. CCTV image quality might often be too poor to extract iris images of sufficient quality but this is merely a hardware question.

We have to accept that our facial and eye images, voice patterns, fingerprints, and DNA etc. are publicly available. It has already been shown, through various experiments that many if not all biometric technologies including iris recognition are susceptible to spoofing attacks: biometric identifiers can be copied and used to create some ▶ artifacts that can deceive several biometric devices available today (▶ Security and Liveness, Overview).

Therefore, the question is not whether biometrics can be copied and forged but rather whether devices can perform accurate liveness testing. The aim of liveness testing is to determine if the biometric data is being captured from a legitimate, live user who is physically present at the point of acquisition. This is especially crucial for remote authentication services performed over open networks where neither the end user's terminal nor the data transmission channels can

be controlled by system operators. An increasing number of vendors are implementing liveness testing into their biometric devices, to guard against the threat of spoofing attacks. However, spoofing-related issues remain unknown to many [2].

Risks of Biometric Spoofing

In order to understand security mechanisms such as liveness detection, it is important to analyze the inherent risks and weaknesses first. Like any other security technology, biometrics also have inherent weaknesses that can potentially lead to security breaches. Susceptibility to spoofing attacks is just one possible weakness inherent to biometric readers (► Biometrics Vulnerabilities: Overview).

Biometric spoofing attacks can be either digital or physical. Digital attacks are defended against by authenticating the biometric reader sending the data and eliminating vulnerable data paths; liveness testing methods are not applicable in this case. Hence this type of vulnerability is not being discussed any further in this article. Physical spoofing of a biometric credential refers to the attack whereby an adversary copies a legitimate biometric to generate a fake artifact and tries to gain access to the system using this artifact. Spoofing attacks may be undertaken with the cooperation of the legitimate user, in an effort to delegate access rights, or without user knowledge by collecting iris pictures from iris recognition systems and infrared cameras, or facial images from camera and surveillance equipment [2].

A biometric system can be used in an access control scenario, either logical or physical, or as a watch list application to detect and identify particular "wanted" individuals. In each case, the purpose of the system and the risks of spoofing are different:

1. In an access control scenario, the system keeps a register of authorized users. An example of such a system is the voluntary, fully-automated immigration control system, IRIS, which operates at several major airports in the UK. In such a scenario, fake iris artifacts could be used to:
   a. Mount attacks against existing enrolments in order to gain unauthorized access – either logical or physical – to the resources protected by the iris recognition system and/or to fraudulently associate an audit trail with an unwitting individual
   b. Enroll into the iris system and then delegate these artifacts across multiple individuals, undermining the integrity of the system
   c. Additionally, a legitimate user could try to repudiate transactions associated with his account or enrolment, claiming instead that they are the result of attacks, due to the inability of the biometric system to ensure liveness
2. In a watch list application, the system keeps a record of people who are being sought by the authorities or who are to be unequivocally denied access to the assets or facilities protected by the biometric device. An example of such a system is the iris-based border control system operated by United Arab Emirates' authorities. In such a setup, it is of course preferable from a registered person's point of view not to be recognized by the system. In connection to the watch lists, an iris artifact could be used to:
   a. A bogus enrolment record can be created so that the wanted/unauthorized person could continue to use the system with his real irises without being detected
   b. A wanted person's iris patterns can be imitated to lead authorities to think (even just temporarily) that someone from the watch list had been caught
   c. A wanted person's iris patterns can be disguised (which had been previously registered in the system) avoid identification and/or gain unauthorized access.

Note that the goals of the ► impostor are different in these two scenarios: for access lists, it's about impersonating a legitimate user, while for the watch lists, it's about disguising one's identity. The latter is believed to be a much easier task, i.e., it is easier to disguise one's iris patterns so that they cannot be recognized anymore than to imitate someone else's iris patterns so closely that a match is achieved.

Spoof-resistance Testing

Governments, academics and an increasing number of industry players are active in the space of testing the resilience of biometric systems against spoofing. While vendors and governments tend to keep their results secret, several test methods and results were published in recent years by the academics and consultants.

In 2002, a report was published by the German Fraunhofer Research Institute, detailing results of much earlier spoofing experiments on face, finger, and iris systems which they carried out in cooperation with the German Federal Institute for Information security (BSI) [3]. This report was among the first to raise a few eyebrows about the security of biometric systems available at that time.

A 2002 issue of c't magazine [4] followed by the description of further attacks. Alongside several other biometric technologies, the authors were able to spoof a low-cost iris recognition device using high-resolution eye images with cut out holes for the pupil which they placed in front of a real eye.

A well-known authority in spoofing, Professor Tsutomu Matsumoto of Yokohama National University in Japan, has published the results of two iris spoofing rounds so far. In 2004, he spoofed three iris recognition cameras with high-resolution photographs with cut out holes for the pupil placed in front of a real eye [5]. Only one of the iris systems did not accept the fake iris for enrolment, but all devices could be spoofed during verification. In 2007, Professor Matsumoto presented another spoofing method using metallic rivets with shiny black rounds heads and printed iris images [6, 7].

Much research has gone into analyzing the effectiveness of spoofing of contact lenses with printed or hand-painted iris patterns as well. In addition to printed photographs and printed/painted contact lenses, other artifacts which could be used to physically spoof iris devices include screen images, video recordings, and artificial eyes (glass, plastic etc.).

Liveness Detection Mechanisms

Biometric experts have been actively researching methods to counter the threat of physical spoofing of biometric samples for more than a decade now. In particular, various liveness detection methods have been conceived and indeed implemented in some devices. However, as every man-made solution can be defeated, efforts are ongoing in this area.

System supervision is the first line of defense against spoofing. The use of several types of spoofing artifacts becomes inconvenient if not impossible, if a human supervisor is present at the point of iris image acquisition. Such examples include the use of photographs or video recordings. Human operators can also detect printed or painted lenses but the costs and inconvenience of such a process make its day-to-day use prohibitive. Additionally, human performance is affected by many factors including tiredness, motivation, sickness etc.

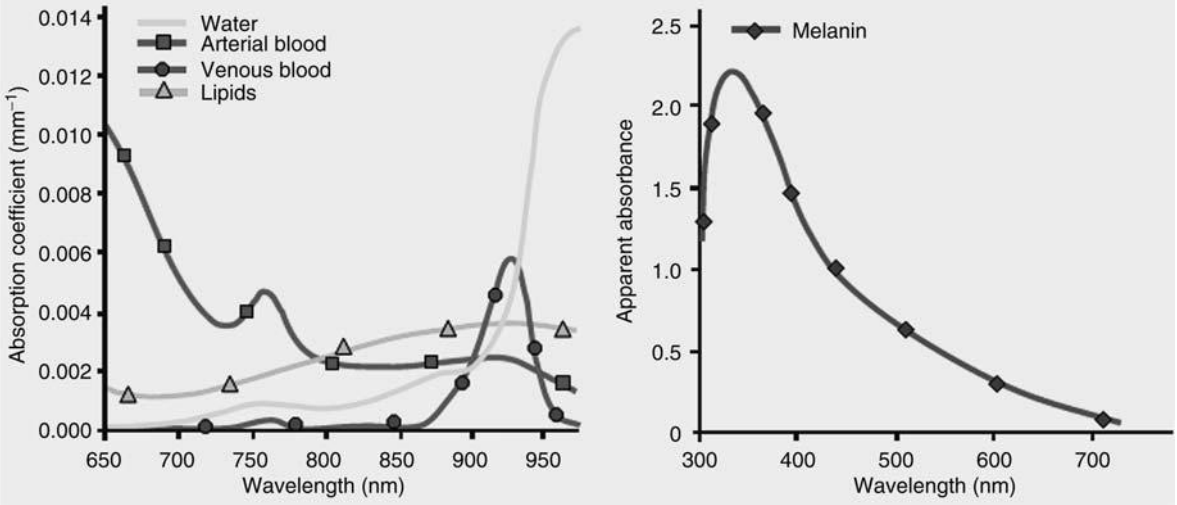
Automated liveness detection can be performed in a biometric device either at the acquisition or processing stage. It usually involves enhancements to software and/or hardware. The presented set of methods is not an exclusive list; however, examples are given for each iris liveness detection category. Automated liveness detection techniques measure and analyze one of the following three characteristics [8].

Intrinsic properties of a living body

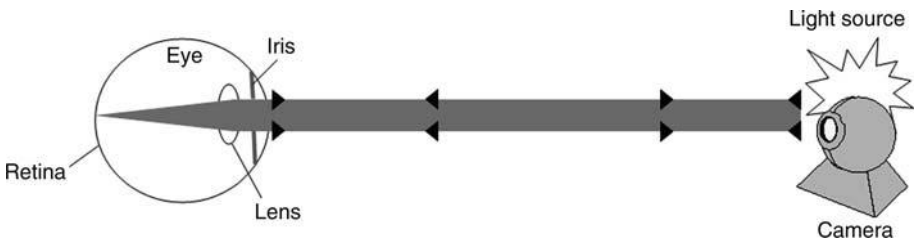
Methods belonging in this category, analyze static characteristics of the material presented to the biometric reader. Such characteristics include density and elasticity (physical properties), capacitance, resistance and permittivity (electrical properties), reflectance and absorbance (spectral properties), color and opacity (visual properties), and chemical content analysis in fluids.

Spectrographic properties of tissue, fat, blood, and pigment can be used to test for liveness in irises. Figure 1 shows that different components of living tissue have distinctive spectrographic signatures. Comparing the fractions of light reflected in 300–1000 nm bandwidth can reveal these spectrographic signatures [9]. If the iris presented to the system is a glass eye, a photograph, or dead tissue, spectrographic analyses could help in detecting the spoofing attack. In fact, ink and paper used to create photographic printouts are often completely ineffective in near-infrared light, which is used during the acquisition of iris images [11].

Retinal light reflections commonly known as the "red-eye effect" can also be used to detect liveness of the eye. Essentially, light entering the eye is reflected back to the light source by the retina; this effect can be captured by a camera. Functional eye cavity optics make the eye to appear red of the pigment (called retinal or visual purple) in the photoreceptors of the retina (Fig. 2). Red-eye effect will occur if the angle between light source, eye and camera is smaller than 2.5° [10].



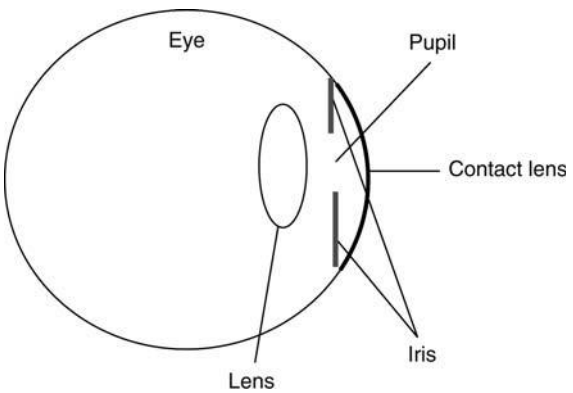
Liveness Detection: Iris. **Figure 1** Light-reflecting properties of different components of living tissue can help to detect iris-spoofing attempts [2, 9, 10].



Liveness Detection: Iris. **Figure 2** The "red-eye effect" – the retina reflects the light entering the eye back to the light source [10].

The iris is a relatively flat internal organ, located behind the cornea and in front of the lens. When a printed or hand-painted contact lens is placed over the eye, the fake "iris" is floating over an external, round surface, the cornea (Fig. 3). Therefore, another intrinsic property for which iris devices can scan is the 3D nature of the layer containing the iris patterns acquired.

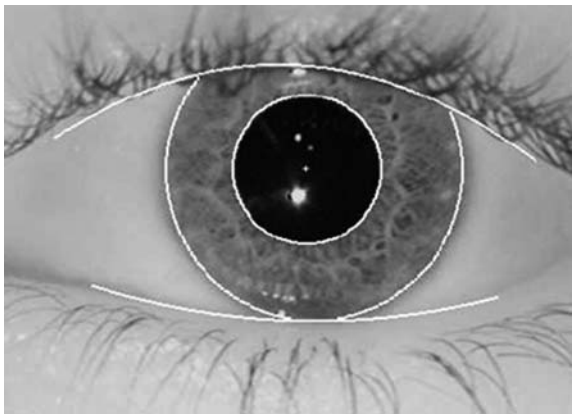
In a natural eye, four optical surfaces reflect light: the front and back surfaces of the cornea as well as the front and back surfaces of the lens (Fig. 4). These reflections are also referred to as Purkinje reflections or images, named after a Czech physiologist. The front surface of the cornea produces the brightest reflection while the back of the lens produces the weakest one. The position of the reflected light determines the position of the reflections – another intrinsic property which can be used to distinguish between natural



Liveness Detection: Iris. **Figure 3** Contact lenses with fake iris patterns float over the curved external surface of the eye whereas the iris is lying in an internal plane inside the eye.

eyes and fake artifacts. A change in the location of the light source should therefore even screen out photographs displaying Purkinje reflections [11]. It might be difficult to capture all four Purkinje reflections at all times due to their varying strength; however, it could be sufficient to analyze the strongest reflections coming from the outer layer of the cornea. Varying positions of near-infrared light diodes used during image acquisition could also be used to analyze this property of the living eye.

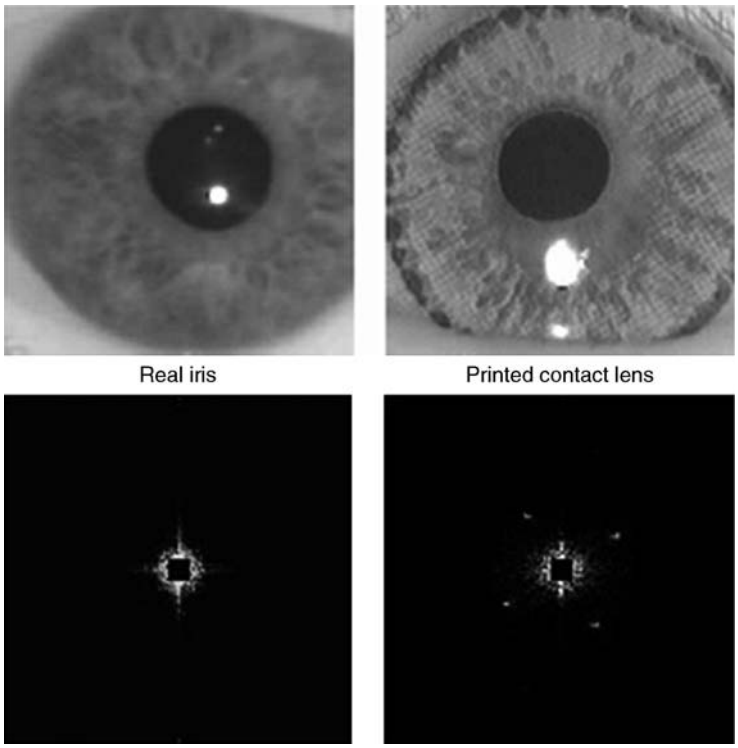
John Daugman, the inventor of iris recognition, has also pointed out that the printing process itself can leave detectable traces on spoofing lenses [11]. A 2D Fourier analysis of the acquired image can show off traces of printing, as demonstrated in Fig. 5. Pacut and Czajka have developed automated methods to analyze artificial frequencies in printed iris images. One great advantage of this method is that, it does not require any additional hardware; it merely analyzes the iris image already captured. However, according to Shanon's theory, the method has a drawback: it fails once the resolution of the printing device is higher than twice the resolution of the analysis camera [12].



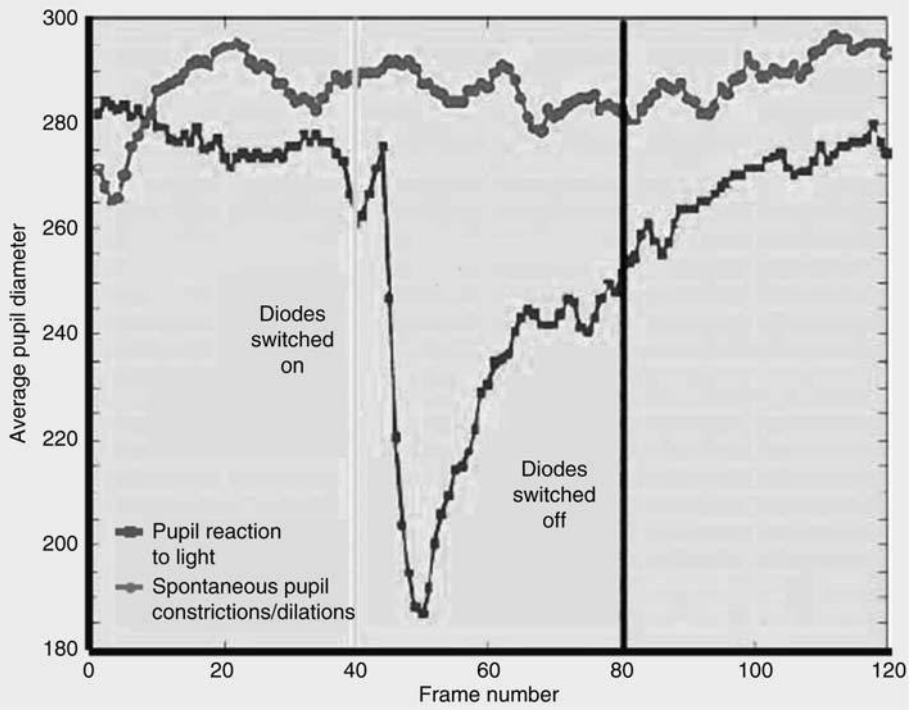
Liveness Detection: Iris. **Figure 4** A picture of a natural eye displaying Purkinje reflections.

Involuntary signals of a living body

Living tissue involuntarily displays dynamic signals which are measurable. These signals can be attributed



Liveness Detection: Iris. **Figure 5** 2D Fourier analysis extracts remnants of the printing process on a contact lens.



Liveness Detection: Iris. **Figure 6** Spontaneous pupil size variations with (square dotted line) and without (round dotted line) any changes in lighting levels [2, 12].

to the nervous system and include pulse, blood pressure, pupillary unrest (hippus), perspiration, blood flow, brain wave signals (EEG), and electrical heart signals (ECG or EKG).

Daugman mentioned the idea of using involuntary signals of the body to measure liveness detection in iris recognition schemes in several of his papers. A very interesting yet little known involuntary signal generated by the human body is the hippus, which is a pupillary steady-state oscillation at about 0.5 Hz, occurring in eyes without any changes in illumination. The coefficient of variation is at least 3% [11], although it declines with advancing age. This liveness detection technique can effectively be used to screen out prosthetic eyes, high-resolution photographs, or dead tissue. The upper graph (with round dots) in Fig. 6 shows involuntary changes in pupil size.

► Iris recognition algorithms need to track the inner and outer boundaries of the iris anyway as part of the extraction process so tracking the changes of pupil size as well as eyelid movements are relatively easy liveness detection methods to implement.

Bodily responses to external stimuli

Finally, it is possible to measure dynamic bodily responses to external stimuli. These liveness detection methods are challenge-response techniques that either look for voluntary (behavioral) responses or involuntary (reflexive) ones.

Behavioral challenge-response methods require user cooperation. As an example, the spoofing resistance of iris recognition products can be enhanced by prompting the user to blink or look left and right, and up and down. If the signal presented to the system is a photograph or video recording, the system is likely to recognize these as fakes.

For iris recognition, an involuntary reflex of the body can be easily triggered by changing illumination levels. The pupil can be driven larger or smaller by changes in lighting conditions, with a response time constant of about 250 ms for constriction and about 400 ms for dilation [11]. The lower graph (with square dots) in Fig. 6 shows the pupillary reflex as a diode is switched on and off.

Another interesting effect which can be observed when the pupil size changes is the nonelastic distortion

of the iris tissue itself. Contact lenses or photographs won't be able to imitate this process.

The Effectiveness of Liveness Testing Methods

Some of the above presented methods have been tested independently with very promising results [3]. However, there is a need for a consistent testing framework to assess the effectiveness of liveness testing methods and market iris products on an ongoing basis. Schemes have been proposed by both the academia and industry [3, 7, 13, 14] but there is still no consensus over an internationally standardized spoof-resistance testing methodology.

The Trade-off between Security and Convenience

Biometric devices should only be spoof-protected to a level corresponding to the nature of operations (i.e., depending on whether operations are mainly security- or convenience-focussed) due to the following limitations of liveness detection methods [2]:

- Firstly, there is a conjecture that for all biometrics, the problem of confirming the vitality of a sample ("liveness testing") is more difficult than to make decisions about matches between templates. The two distributions of similarity generated by "genuine" and "spoof" samples for the same person are likely to be closer and to have more overlap in the vitality test than the two distributions that are generated in a template matching test by "same" and "different" persons without any spoofing effort [9]. In other words, liveness testing can adversely affect recognition performance ("Security and Liveness, Overview").
- Secondly, liveness tests have the propensity to increase the time to acquire the biometric sample, thus reducing user convenience.
- Finally, the incorporation of liveness tests into a device usually also means increasing hardware/software costs.

Summary

Spoofing is a real concern with regards to the security of biometric systems. More and more successful spoofing attempts are being published and even though the sophistication of these attacks is on the rise, spoofing is still in its infancy. In particular, contact lenses with hand-painted and printed iris patterns are expected

Liveness Detection: Iris. Table 1 Overview of discussed liveness detection methods for iris recognition with an indication of their effectiveness

Category	Countermeasure	Targeted artifacts
Extraction of intrinsic properties (static)	Spectrographic analysis	All
	Near-infrared illumination	All involving inks and dyes
	Red-eye effect	All except patterned contact lenses
	3D curvature of iris surface	Patterned contact lenses
	Purkinje reflections	All except patterned contact lenses
	Frequency spectrum analysis	Printed artifacts
Analysis of involuntary signals (dynamic)	Pupillary unrest (hippus)	All except patterned contact lenses which only party cover the real iris
	Eyelid movements	All except patterned contact lenses
Challenge-response methods (dynamic)	Eye movements (blinking, looking in various directions)	All except patterned contact lenses
	Pupillary light reflex	All except patterned contact lenses which only party cover the real iris
	Nonelastic distortion of iris tissue	All

to pose an increasing threat due to enhancements of ink quality and printing technologies. Furthermore, patterned lenses are relatively difficult to detect when compared to some of the other spoofing methods. Both the industry and academia are focusing their efforts to make biometric devices more robust but every countermeasure can eventually be circumvented. Thus research and development efforts must be ongoing.

This article illustrates that it is possible to combat physical spoofing attacks with liveness testing (Table 1) but all of these countermeasures come at a certain price, often affecting user convenience, system prices, or matching accuracy. Therefore, it is crucial to select a device that incorporates spoofing countermeasures to a level of sophistication and effectiveness that matches the requirements of the application.

As spoofing techniques are swiftly evolving and countermeasures have only a limited life cycle, in addition to the necessary research and development efforts it is of great importance to perform standardized, vendor-independent tests of robustness and to assess on a regular basis the overall level of security provided by biometric systems.

Related Entries

- ▶ Anti-spoofing
- ▶ Liveness Detection
- ▶ Spoofing; Biometric Vulnerabilities

References

1. Toth, B.: Biometric ID card debates. "Biometrie" Newsletter of ISACA Chapter Switzerland, Germany and Austria (2005)
2. Toth, B.: Biometric Liveness Detection. Information Security Bulletin, ISB1008, CHI Publishing 291–297 (2005)
3. Czajka, A., Strzelczyk, P., Pacut, A.: Making iris recognition more reliable and spoof resistant. SPIE Newsroom, doi: 10.1117/2.1200706.0614 (2007)
4. Thalheim, L., Krissler, J., Ziegler, P-M.: Body Check. Biometric Access Protection Devices and their Programs Put to the Test, c't Magazine (2002)
5. Matsumoto, T.: Artificial fingers and irises: Importance of vulnerability analysis. In: Proceedings of the 7th International Biometrics Conference, London. Elsevier (2004)
6. Matsumoto, T., Kusuda, T.: IECIE Technical Report on Biometrics Security (2007)
7. Matsumoto, T.: Assessing the security of advanced biometric systems: Finger, vein and iris. In: Proceedings of the 10th International Biometrics Conference, London. Elsevier (2007)
8. Woodward, J.D., Orlans, J.M., Higgins, P.T.: Biometrics. Identity Assurance in the Information Age, Osborne McGraw-Hill Osborne Media Berkeley, California, USA (2003)
9. Daugman, J.G.: Iris recognition and anti-spoofing countermeasures. In: Proceedings of the 7th International Biometrics Conference, London. Elsevier (2004)
10. Toth, B.: Liveness detection for iris recognition, In: Proceedings of Biometrics and E-Authentication over Open Networks, NIST Gaithersburg (2005)
11. Daugman, J.G.: Recognizing persons by their iris patterns: Countermeasures against subterfuge. In: Jain, A.K., et al. (eds.) Biometrics: Personal Identification in a Networked Society, Kluwer Academic Publishers, Dordrecht, The Netherlands(1999)
12. Pacut, A., Czajka, A.: Liveness detection for iris biometrics. In: Proceedings of 40th IEEE International Carnahan Conference on Security Technology, Lexington, KY. IEEE (2006)
13. Dunstone, T.: Current Status of Biometrics in Australia and New-Zealand including Vulnerability Assessments. In: Proceedings of the 10th International Biometrics Conference, London. Elsevier (2007)
14. Nanavati, S.: Spoofing Biometrics: Deploying Imperfect Biometrics. In: Proceedings of the 10th International Biometrics Conference, London. Elsevier (2007)

Live-Scan Furrow Device

It refers to a device, able to read the ridge-valley pattern present on finger tips, palms, and foot soles. It can be considered as a generic name grouping all the fingerprint, palmprint, handprint, and soleprint devices.

▶ Fingerprint, Palmprint, Handprint and Soleprint Sensor

Live-Scan Sensor

A live-scan sensor is a sensor that allows to capture and digitize biometric data in real time. As opposed to live-scan acquisition, in off-line acquisition, data is not digitized in real time (e.g., a fingerprint image is first obtained by smearing ink on the fingertip and creating

an inked impression on paper, and the inked impression is digitized by scanning the paper).

► Fingerprint Databases and Evaluation

Local Adaptive Thresholding

Local adaptive thresholding is used to convert an image consisting of gray scale pixels to just black and white scale pixels. Usually a pixel value of 0 represents white and the value 255 represents black with the numbers from 1 to 254 representing different gray levels. Unlike the global thresholding technique, local adaptive thresholding chooses different threshold values for every pixel in the image based on an analysis of its neighboring pixels. This is to allow images with varying contrast levels where a global thresholding technique will not work satisfactorily. There are a number of different forms of adaptive thresholding algorithm reported in the image processing literature.

► Hand Vein

Local Descriptors

► Local Image Features

Local Fusion

Local fusion in the framework or multi-biometric score fusion refers to user-specific score fusion techniques in which each fusion function is trained based exclusively on data associated with the claimed user (both genuine and impostor scores).

► Fusion, User-Specific

Local Image Features

Krystian Mikolajczyk[1], Tinne Tuytelaars[2]
[1]School of Electronics and Physical Sciences, University of Surrey, Guildford, Surrey, UK
[2]Department of Electrical Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, Leuven, Belgium

Synonyms

Interest points; Keypoints; Local descriptors

Definition

A ► local feature is an image pattern which differs from its immediate neighborhood. It is usually associated with a change of an image property or several properties simultaneously, though it is not necessarily localized exactly on this change. The image properties commonly considered are intensity, color, and texture. Figure 1 shows some examples of local features in a contour image (left) as well as in a grayvalue image (right). Local features can be points, but also edgels or small image patches. Typically, some measurements are taken from a ► region centered on a local feature and converted into descriptors. The descriptors can then be used for various applications. Three broad categories of feature ► detectors can be distinguished based on their possible usage. It is not exhaustive or the only way of categorizing the features but it emphasizes different properties required by the usage scenarios. First, one might be interested in a specific type of local features, as they may have a specific semantic interpretation in the limited context of a certain application. For instance, edges detected in aerial images often correspond to roads; blob detection can be used to identify impurities in some inspection task; etc. These were the first applications for which local feature detectors have been proposed. Second, one might be interested in local features since they provide a limited set of well localized and individually identifiable anchor points. What the features actually represent is not really relevant, as long as their location can be determined accurately and in a stable manner over time. This is for instance the situation in most matching or tracking applications, and especially for camera

calibration or 3D reconstruction. Other application domains include pose estimation, image alignment, or mosaicing. A typical example here is the features used in the KLT tracker [1]. Finally, a set of local features can be used as a robust image representation, that allows to recognize objects or scenes without the need for segmentation. Here again, it does not really matter what the features actually represent. They do not even have to be localized precisely, since the goal is not to match them on an individual basis, but rather to analyze their statistics. This way of exploiting local features was first reported in the seminal work of [2] and soon became very popular, especially in the context of object recognition (both for specific objects as well as for category-level recognition). Other application domains include scene classification, texture analysis, image retrieval, and video mining.

Introduction

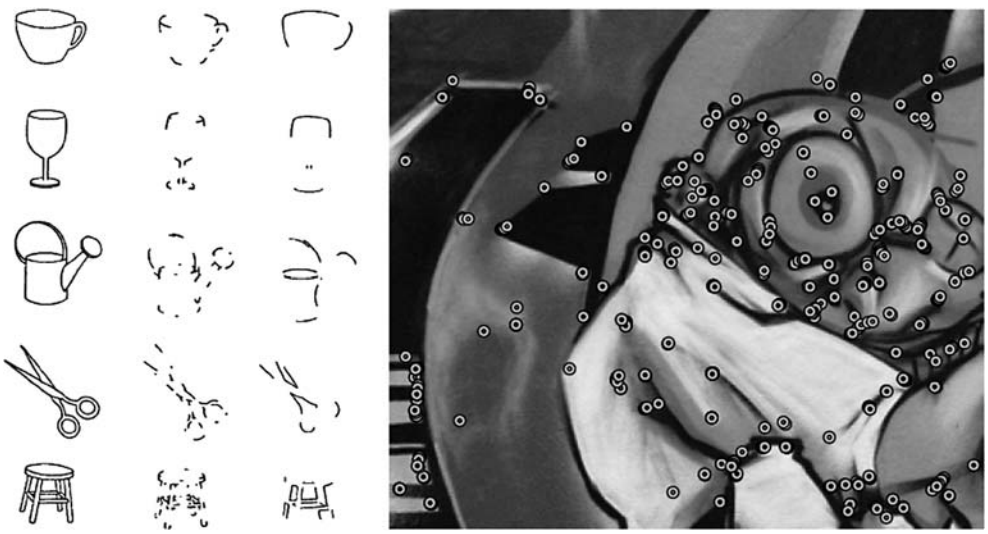
The first publication on local features appeared after the observation on the importance of corners and junctions in visual recognition [3] (see Fig. 1). Since then a large number of algorithms have been suggested for extracting ▶ interest points at the extrema of various functions computed on the digital shape. Also, it has been understood early on in the image processing and visual pattern recognition field that intersections of straight lines and straight corners are strong indications of man made structures. Such features have been used in the first series of applications from line drawing images [4] and photomosaics [5]. First monographs on digital image processing [6, 7] and later editions served to establish the field on a sound theoretical foundation. Several survey articles on local features appeared recently [8–10].

Interest points are now the preferred strategy for solving a wide variety of problems, from wide baseline matching and the recognition of specific objects to the recognition of object classes. Additionally, similar ideas have been applied to texture recognition, scene classification, robot navigation, visual data mining, and symmetry detection, to name just a few application domains.

Local ▶ invariant features not only allow to find correspondences, in spite of large changes in viewing conditions, occlusions, and image clutter (wide baseline matching), but also yield an interesting description of the image content for image retrieval and object or scene recognition tasks (both for specific objects as well as categories). To put this into context, some alternative strategies are briefly summarized, including global features, image segments, and exhaustive and random sampling of features.

Global Features

In the field of image retrieval, many global features have been proposed to describe the image content, with color histograms and variations thereof as a



Local Image Features. **Figure 1** Illustration of local features in line drawing images and a grayvalue image.

typical example [11]. This approach works surprisingly well, at least for images with distinctive colors, as long as it is the overall composition of the image as a whole that the user is interested in, rather than the foreground object. Indeed, global features cannot distinguish foreground from background, and mix information from both parts together.

Global features have also been used for object recognition, resulting in the first appearance-based approaches to tackle this challenging problem. [12] and later [13] proposed to compute a principal component analysis of a set of model images and to use the projections onto the first few principal components as descriptors. Compared to the purely geometry-based approaches tried before, the results of the novel appearance-based approach were striking. A whole new range of natural objects could suddenly be recognized. However, being based on a global description, image clutter and occlusions again form a major problem, limiting the usefulness of the system to cases with clean backgrounds or where the object can be segmented out, e.g., relying on motion information.

Image Segments

An approach to overcome the limitations of the global features is to segment the image in a limited number of regions or segments, with each such region corresponding to a single object or a part thereof. However, this raises a chicken-and-egg problem as image segmentation is a very challenging problem in itself, which in general requires a high-level understanding of the image content. For generic objects, color and texture cues are insufficient to obtain meaningful segmentations.

Sampled Features

A way to deal with the problems encountered with global features or image segmentation is to *exhaustively sample* different subparts of the image at each location and scale. For each such image subpart, global features can then be computed. This approach is also referred to as a *sliding window* based approach. It has been especially popular in the context of face detection, but has also been applied for the recognition of specific objects or particular object classes such as pedestrians or cars.

By focusing on subparts of the image, these methods are able to find similarities between the queries and the models in spite of changing backgrounds, even if the object covers only a small percentage of the total image area. In the bottom, they still do not manage to cope with partial occlusions, and the allowed shape variability is smaller than what is feasible with a local feature based approach. However, by far the biggest drawback is the inefficiency of this approach. Each and every subpart of the image must be analyzed, resulting in thousands or even millions of features per image. This requires extremely efficient methods which significantly limits the scope of possible applications. To overcome the complexity problems sparser *fixed grid sampling* of image patches can be used. It is however difficult to achieve invariance to geometric deformations for such features. The approach can tolerate some deformations due to dense sampling over possible locations, scales, poses etc. but the individual features are not invariant. As a result, sampled features cannot be used when the goal is to find precise correspondences between images. However, for some applications such as scene classification or texture recognition, they may well be sufficient.

In a similar vein, rather than using a fixed grid of patches, a *random sampling* of image patches can also be used. This gives a larger flexibility in the number of patches, the range of scales or shapes, and their spatial distribution. Random patches are in fact a subset of the dense patches, and are used mostly to reduce the complexity. Their repeatability is poor hence they work better as an addition to the regular features rather than as a stand alone method.

Finally, to overcome the complexity problems while still providing a large number of features with better than random localization one can sample features uniformly from edges. This proved useful for dealing with wiry objects well represented by edges and curves.

Properties of the Ideal Local Feature

Local features typically have a spatial extent, i.e., the local neighborhood of pixels mentioned above. In contrast to classical segmentation, this can be any subset of an image. The region boundaries do not have to correspond to the changes in image appearance such as color or texture. Also, multiple regions may overlap,

and "uninteresting" parts of the image such as homogeneous areas can remain uncovered.

Ideally, one would like such local features to correspond to semantically meaningful object parts. In practice, however, this is unfeasible, as this would require high-level interpretation of the scene content, which is not available at this early stage. Instead, detectors select local features directly based on the underlying intensity patterns.

Good features should have the following properties:

- *Repeatability*: Given two images of the same object or scene, taken under different viewing conditions, a high percentage of the features detected on the scene part visible in both images should be found in both images.
- *Distinctiveness/informativeness*: The intensity patterns underlying the detected features should show a lot of variation, such that features can be distinguished and matched.
- *Locality*: The features should be local, so as to reduce the probability of occlusion and to allow simple model approximations of the geometric and photometric deformations between two images taken under different viewing conditions (e.g., based on a local planarity assumption).
- *Quantity*: The number of detected features should be sufficiently large, such that a reasonable number of features are detected even on small objects. However, the optimal number of features depends on the application. Ideally, the number of detected features should be controllable over a large range by a simple and intuitive threshold. The density of features should reflect the information content of the image to provide a compact image representation.
- *Accuracy*: The detected features should be accurately localized, in both image location, with respect to scale and possibly shape.
- *Efficiency*: Preferably, the detection of features in a new image should allow for time-critical applications.

  Repeatability, arguably the most important property of all, can be achieved in two different ways: either by invariance or by robustness.
- *Invariance*: When large deformations are to be expected, the preferred approach is to model these mathematically if possible, and then develop methods for feature detection that are unaffected by these mathematical transformations.
- *Robustness*: In case of relatively small deformations, it often suffices to make feature detection methods less sensitive to such deformations, i.e., the accuracy of the detection may decrease, but not drastically. Typical deformations that are tackled using robustness are image noise, discretization effects, compression artifacts, blur, etc. Also geometric and photometric deviations from the mathematical model used to obtain invariance are often overcome by including more robustness.

Clearly, the importance of these different properties depends on the actual application and settings, and compromises need to be made.

Repeatability is required in all application scenarios and it directly depends on the other properties like invariance, robustness, quantity etc. Depending on the application, increasing or decreasing them may result in higher repeatability.

Distinctiveness and locality are competing properties and cannot be fulfilled simultaneously: the more local a feature, the less information is available in the underlying intensity pattern and the harder it becomes to match it correctly, especially in database applications where there are many candidate features to match to. On the other hand, in case of planar objects and/or purely rotating cameras (e.g., in image mosaicing applications), images are related by a global homography, and there are no problems with occlusions or depth discontinuities. Under these conditions, the size of the local features can be increased without problems, resulting in a higher distinctiveness.

Similarly, an increased level of *invariance* typically leads to a reduced *distinctiveness*, as some of the image measurements are used to lift the degrees of freedom of the transformation. A similar rule holds for *robustness versus distinctiveness*, as typically some information is disregarded (considered as noise) to achieve robustness. As a result, it is important to have a clear idea on the required level of invariance or robustness for a given application. It is hard to achieve high invariance and robustness at the same time and invariance, which is not adapted to the application, may have a negative impact on the results.

Accuracy is especially important in wide baseline matching, registration, and structure from motion applications, where precise correspondences are

needed to, e.g., estimate the epipolar geometry or to calibrate the camera setup.

Quantity is particularly useful in some class-level object or scene recognition methods, where it is vital to densely cover the object of interest. On the other hand, a high number of features have in most cases a negative impact on the computation time and it should be kept within limits. Also robustness is essential for object class recognition, as it is impossible to model the intra-class variations mathematically, so full invariance is impossible. For these applications, an accurate localization is less important. The effect of inaccurate localization of a feature detector can be countered, up to some point, by having an extra robust descriptor, which yields a feature vector that is not affected by small localization errors.

Related Entries

- ► Gabor filter
- ► Image descriptors
- ► Local binary pattern
- ► Local Feature Filters
- ► Matching
- ► Registration

References

1. Shi, J., Tomasi, C.: "Good features to track," in: Proceedings of the Conference on Computer Vision and Pattern Recognition, pp. 593–600 (1994)
2. Schmid, C., Mohr, R.: Local gray-value invariants for image retrieval. IEEE Trans. Pattern Anal. Mach. Intell. **19**(5), 530–534 (1997)
3. Attneave, F.: Some informational aspects of visual perception. Psychol. Rev. **61**, 183–193 (1954)
4. Freeman, H.: A review of relevant problems in the processing of line-drawing data, AII, pp. 155–174 (1969)
5. Milgram, D.: Computer methods for creating photomosaics. IEEE Trans. Comput. **23**, 1113–1119 (1975)
6. Duda, R., Hart, P.: Pattern Classification and Scene Analysis. Wiley, New York (1973)
7. Rosenfeld, A.: Picture processing by computer. ACM Comput. Surv. **1**(3), 147–176 (1969)
8. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. (2005)
9. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Gool, L.V.: A comparison of affine region detectors. Int. J. Comput. Vis. (2005)
10. Tuytelaars, T., Mikolajczyk, K.: Local Invariant Feature Detectors: A Survey, Foundations and Trends in Computer Graphics and Vision (2008)
11. Swain M., Ballard, D.: Color indexing. Int. J. Comput. Vis. **7**(1), 11–32 (1991)
12. Turk, M.A., Pentland, A.P.: Eigenfaces for face recognition, in *The Conference on Computer Vision and Pattern Recognition*, pp. 586–591 (1991)
13. Murase, H., Nayar, S.: Visual learning and recognition of 3D objects from appearance. Int. J. Comput. Vis. **14**(1), 5–24 (1995)

Local Image Filters

Abdenour Hadid, Matti Pietikäinen
Machine Vision Group, Department of Electrical and Information Engineering, University of Oulu, Finland

Synonyms

Gabor features; LBP features

Definition

Local feature filters can be defined as operators (or filters) which are applied to an image in order to extract local characteristics describing (some) important information in the image. For instance, these characteristics (or ► features) can be used to detect, recognize, and analyze the objects in the image. They can also facilitate the interpretation or further processing of the image. In contrast to global features which describe the overall content and shape of the objects in the image, local features define specific information in local regions. Among the most effective operators for feature extraction are Gabor filter and local binary pattern (LBP). Gabor filters are linear bandpass filters computed for images at different orientations and scales. The impulse response of a Gabor filter is defined by a harmonic function multiplied by a Gaussian function. Local binary pattern is a nonlinear operator which labels the pixels of an image by thresholding the neighborhood of each pixel with the value of the center pixel and considers the result as a binary number. LBP labels can be regarded as local primitives such as curved edges, spots, flat areas, etc. The histogram of

the labels can be then used as a feature vector (or image representation). Gabor filtering and LBP operator are powerful means of analyzing biometric data such faces, irises, fingerprints, palmprints, etc.

Introduction

Typically, biometric systems operate by acquiring biometric data from which features are extracted and matched against those of the templates which are stored in the database [1]. This involves two crucial aspects: feature extraction and classifier design. The aim of feature extraction is to find good descriptors which are easy to compute and have high ► extra-class variance (i.e., between different persons) and low ► intra-class variance, which means that the descriptor should be robust with respect to aging of the subjects, alternating illumination and other factors. Obviously, if inadequate features are adopted, even the most sophisticated classifiers (i.e., comparison schemes) will fail to accomplish the given recognition task. Therefore, feature extraction is a very important task in any biometric system.

Different global (or holistic) methods such as Principal Component Analysis (PCA) have been widely studied and applied to biometrics but lately local features have gained more attention due to their robustness to challenges such as pose and illumination changes. In this context, feature extraction using Gabor filtering or LBP has gained increasing attention in various biometric applications. A notable example is iris recognition, in which approaches based on multichannel Gabor filtering have been highly successful. Gabor filters have also been widely used, e.g., in fingerprint [2] and palmprint analysis [3]. Also, the well-known Elastic Bunch Graph Matching (EBGM) method is based on Gabor filter responses at certain fiducial points to recognize faces [4]. More recently, LBP features have provided excellent results in various biometric applications [5, 6, 7, 8]. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze biometric data in challenging real-time settings.

The Gabor and LBP methods provide complementary information for analysis: LBP captures small and fine details (or micro features) while Gabor filters encode appearance information over a broader range of scales (macro features).

Gabor Filters

The theory behind Gabor filters started from the original work of Dennis Gabor who proposed in 1946 to represent signals as a combination of elementary functions [9]. Those particular elementary functions are now known as Gabor elementary functions (GEF). However, the use of Gabor filters in image processing started from the work of Granlund who extended the elementary 1-d Gabor functions to 2-d elementary functions and used them in the development of a general picture processing operator [10]. Later, Daugman proposed the generalization of 1-d gabor functions to two dimensions and importantly showed the equivalence between a structure based on the 2-d Gabor functions and the organization and the characteristics of the mammalian visual system [11]. These physiological findings are undoubtedly behind the great impact of Gabor research especially in image processing.

Roughly speaking, Gabor filtering in image processing consists of applying a set of 2-d Gabor elementary functions of various parameters (e.g., different dilations and rotations) to an input image thus obtaining Gabor image features (i.e., feature space). These extracted Gabor features can be then used directly as feature vectors for analysis (e.g., biometric recognition) or can first be transformed into new feature vectors (e.g., [8]). So, typically, an input image $I(x, y)$ is convolved with a 2-d Gabor function $g(x,y)$ to obtain a Gabor feature image $r(x, y)$ as follows:

$$r(x,y) = I(x,y) * g(x,y) = \int \int_{-\infty}^{+\infty} I(d\xi, d\eta) g(x - d\xi, y - d\eta)\ d\xi\ d\eta.$$

There are several forms of Gabor elementary functions (GEFs) which can be designed to be highly selective in frequency while displaying good spatial localization. GEFs can be also seen as bandpass filters which can be configured to be used for feature extraction. A GEF is defined as a Gaussian modulated by a sinusoid (cosine function) as follows:

$$g_{\lambda,\theta,\psi,\sigma,\gamma}(x,y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \psi\right),$$

where $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$, λ represents the wavelength of the cosine factor, θ represents the orientation of the normal to the parallel stripes of a Gabor function, ψ is the phase offset, σ refers to the variance of the Gaussian function and γ is the spatial aspect ratio and specifies the ellipticity of the support of the Gabor function. Figure 1 shows an example of a typical 2-d Gabor filter.

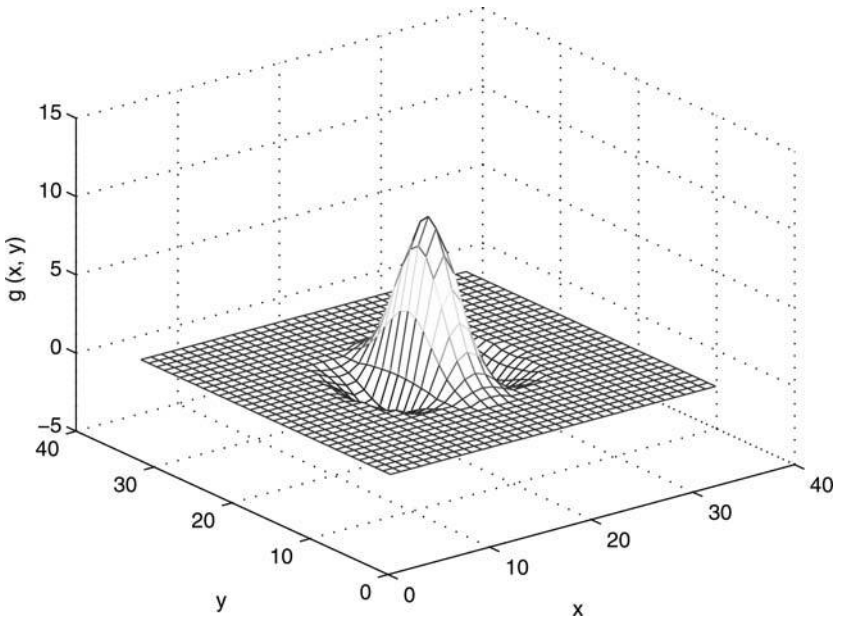
Local Binary Patterns

The LBP texture analysis operator [12], introduced by Ojala et al., is defined as a gray-scale invariant texture measure, derived from a general definition of texture in a local neighborhood. It is a powerful means of texture description and among its properties in real-world applications are its discriminative power, computational simplicity, and tolerance against monotonic gray-scale changes. The original LBP operator forms labels for the image pixels by thresholding the 3×3 neighborhood of each pixel with the center value and considering the result as a binary number. The histogram of these $2^8 = 256$ different labels can then be used as a texture descriptor.

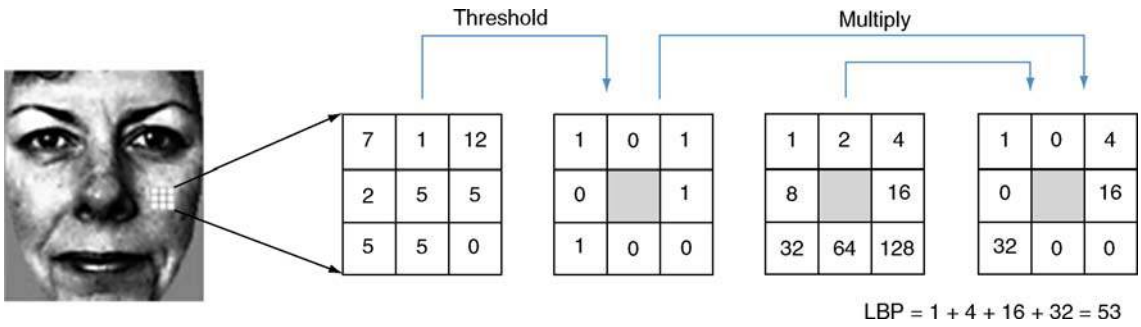
The operator has been extended to use neigborhoods of different sizes [13]. Using a circular neighborhood and bilinearly interpolating values at noninteger pixel coordinates allows any radius and number of pixels in the neighborhood. The notation (P, R) is generally used for pixel neighborhoods to refer to P sampling points on a circle of radius R. The calculation of the LBP codes can be easily done in a single scan through the image. See Fig. 2 for an illustration of the basic LBP operator. The value of the LBP code of a pixel (x_c, y_c) is given by $\mathrm{LBP}_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p$, where g_c corresponds to the gray value of the center pixel (x_c, y_c), g_p refers to gray values of P equally spaced pixels on a circle of radius R, and s defines a thresholding function as follows:

$$s(x) = \begin{cases} 1, \textit{if } x \geq 0; \\ 0, \textit{otherwise}. \end{cases}$$

Another extension to the original operator is the definition of so called *uniform patterns* [13]. This extension was inspired by the fact that some binary patterns occur more commonly in texture images than in others. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is traversed circularly. For example, the patterns 00000000 (0 transitions), 01110000 (2 transitions),



Local Image Filters. Figure 1 An example of a typical 2-d Gabor filter.



Local Image Filters. **Figure 2** The basic LBP operator.

and 11001111 (2 transitions) are uniform whereas the patterns 11001001 (4 transitions) and 01010011 (6 transitions) are not. In the computation of the LBP labels, uniform patterns are used so that there is a separate label for each uniform pattern and all the nonuniform patterns are labeled with a single label. For example, when using (8, R) neighborhood, there are a total of 256 patterns, 58 of which are uniform, which yields 59 different labels. This yields the following notation for the LBP operator: $\mathrm{LBP}^{U2}_{P,R}$. The subscript represents using the operator in a (P, R) neighborhood. Superscript $u2$ stands for using only uniform patterns and labeling all remaining patterns with a single label. Each bin (LBP label) can be regarded as a microtexton. Local primitives which are codified by these bins include different types of curved edges, spots, flat areas, etc.

The original LBP operator was defined to only deal with spatial information. Recently, it has been extended to a spatiotemporal representation for dynamic texture analysis. This has yielded the so called Volume Local Binary Pattern operator (VLBP) [6]. The idea behind VLBP consists of looking at dynamic texture as a set of volumes in the (X,Y, T) space where X and Y denote the spatial coordinates and T denotes the frame index (time). The neighborhood of each pixel is thus defined in three dimensional space. Then, similar to LBP in spatial domain, volume textons can be defined and extracted into histograms. Therefore, VLBP combines motion and appearance together to describe dynamic texture. Later, to make the VLBP computationally simple and easy to extend, the cooccurrences of the LBP on three orthogonal planes (LBP-TOP) were also introduced [6]. LBP-TOP consists then in considering three orthogonal planes: XY, XT, and YT, and concatenating local binary pattern cooccurrence statistics in these three directions. The circular neighborhoods are generalized to elliptical sampling to fit to the space-time statistics.

In the LBP approach to texture classification [13], the occurrences of the LBP codes in an image are collected into a histogram. The classification is then performed by computing simple histogram similarities. However, considering a similar approach for biometric (e.g., facial) image representation results in a loss of spatial information and therefore one should codify the texture information while retaining also their locations. One way to achieve this goal using texture operators is to build several local descriptions of the face and combine them into a global description [5, 8]. Figure 3 shows an example of an LBP based facial representation. Such local descriptor based methods have been gaining interest lately which is understandable given the limitations of the holistic representations. These local feature based methods seem to be more robust against variations in pose or illumination than holistic methods.

Applications

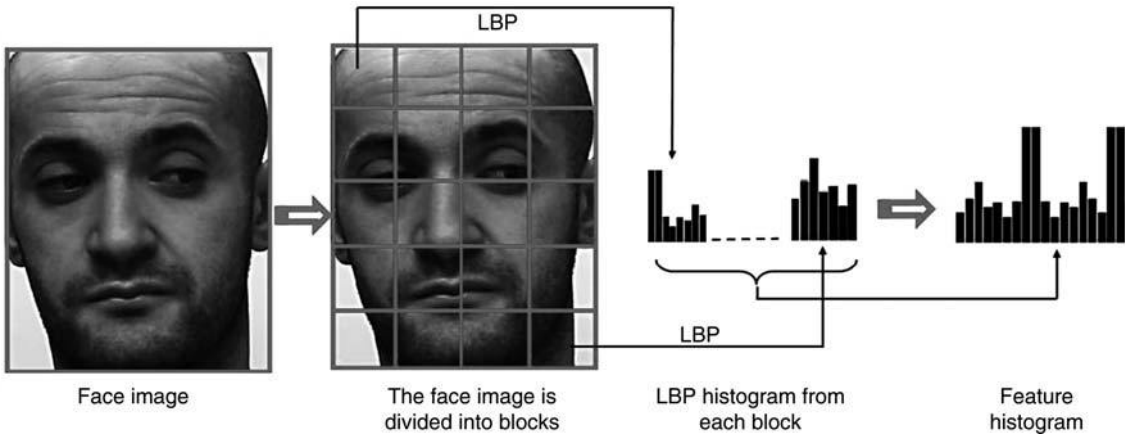
There is a considerable amount of research concerning Gabor filtering based biometric recognition, especially in iris, face, fingerprint, and plamprint recognition [1, 14, 2, 15, 4, 3]. For instance, Daugman developed the pioneering approach to iris recognition [14]. He used 2-d Gabor filtering to extract local texture features from iris images, resulting in an IrisCode representation with 2,048 bits. Then, the recognition was done simply by computing the Hamming distance between a pair of iris representations. The system was tested on billions of iris images and the results showed excellent recognition rates very close to 100% (only one false acceptance in 151,000 imposter tests for one

false rejection in 128,000 tests). Nowadays, many commercial iris-recognition systems are based on Daugman's algorithms.

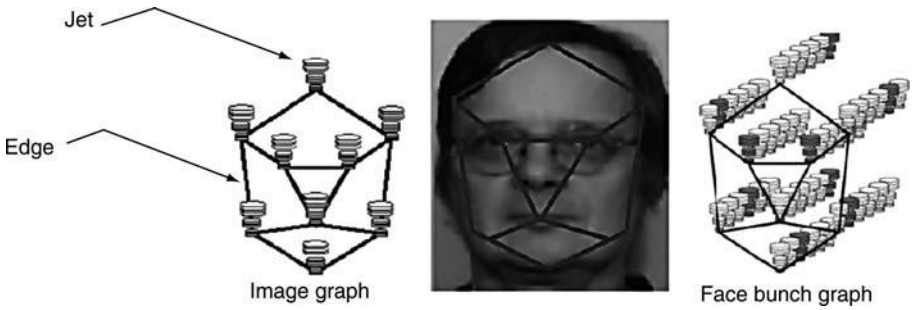
Gabor filter based features have also been widely applied to fingerprint recognition systems. A notable example is the system developed by Jain [2] using a bank of Gabor filters to capture both the global pattern of ridges and valleys and the local characteristics in fingerprint patterns yielding feature vectors called FingerCodes. The fingerprint matching is based on the Euclidean distance between two corresponding FingerCodes. The system achieved remarkable recognition rates.

In face recognition, Lades et al. developed a Gabor based system using dynamic link architecture (DLA) framework which recognizes faces by extracting Gabor jets at each node of a rectangular grid over the face image [15]. Later, Wiskott et al. extended the approach and developed the well-known Gabor wavelet-based elastic bunch graph matching (EBGM) method to label and recognize faces [4]. In the EBGM algorithm, faces are represented as graphs with nodes positioned at fiducial points (such as the eyes, the tip of the nose, etc.) and edges labeled with distance vectors (see Fig. 4). Each node contains a set of Gabor wavelet coefficients, known as a jet. Thus, the geometry of the face is encoded by the edges while the gray value distribution (texture) is encoded by the jets. The identification of a new face consists of determining among the constructed graphs the one which maximizes the graph similarity function.

In palmprint recognition, for instance, Zhang et al. [3] proposed the use of 2-d Gabor phase encoding scheme for palmprint feature extraction and representation. First, the central areas from palmprint images are segmented and then 2-d Gabor based features are extracted. The recognition is performed using normalized Hamming distance. Experiments on a database of 7,752 low resolution palmprint images showed good performance. It is worth nothing that the use of Gabor filtering in biometrics is not limited to the examples cited earlier as many other Gabor based biometric



Local Image Filters. **Figure 3** Example of an LBP based facial representation.



Local Image Filters. **Figure 4** An example of facial representation with the elastic bunch graph principle (EBGM).

systems for commercial applications have been successfully developed. In addition, Gabor features have also been successfully used with other biometric modalities including speech, gait, ear, etc.

LBP has also been successfully used in various biometric applications such as face, activity, iris, and palmprint recognition. The most remarkable application of LBP in biometrics is face analysis [5]. The idea consists in dividing the faces into several regions (or blocks) from which the local binary pattern histograms are computed and concatenated into a single, spatially enhanced feature histogram (Fig. 3). In such a representation, the texture of facial regions is encoded by the LBP while the shape of the face is recovered by the concatenation of different local histograms. The LBP methodology has attained an established position in face analysis research and several research groups around the world have adopted similar approach to different tasks such as near-infrared based face recognition, gender recognition, head pose estimation and 3D face recognition. A bibliography of LBP-related research in facial image analysis can be found at http://www.ee.oulu.fi/research/imag/texture/lbp/bibliography/.

The spatiotemporal versions of LBP (VLBP and LBP-TOP) have also been successfully applied to person analysis and identification from video sequences, including face, facial expression, visual speech and activity recognition. To recognize six prototypic emotions (anger, disgust, fear, joy, sadness, and surprise) from videos, Zhao et al. [6] divided the face sequences into several overlapping block volumes, extracted LBP-TOP (or VLBP) histograms from each block and then concatenated them to obtain a single histogram representing the appearance and motion of the facial expression in the face sequences. This approach does not require error-prone segmentation of lips and other facial features and it is robust against monotonic gray scale changes caused, for example, by illumination and skin color variations, and errors in face alignment. Hadid et al. [7] also adopted spatiotemporal LBP for face recognition from videos with excellent results. Starting from the observation that VLBP features consist of both intra and extra personal information (corresponding to both facial expression and identity), they proposed a robust recognition system using VLBP with AdaBoost learning. The idea was to classify the VLBP facial information into intra and extra classes, and then use only the ► extra-class VLBP features for recognition. This was achieved by looking at a face sequence as a selected set of rectangular prisms (volumes) from which local histograms of extended VLBP code occurrences are extracted. Then, a boosting approach is used for selecting only the most discriminative spatiotemporal patterns for face recognition while discarding the patterns which may hinder the recognition process.

There are many other very successful biometric applications based on Gabor filtering or LBP features. The approaches based on Gabor filtering or wavelets measuring the frequency contents of image points or regions at different resolutions and orientations encode appearance information over a broad range of scales (macro features) while LBP operator captures smaller and finer details (or micro features). This makes Gabor filtering and LBP operator powerful means for extracting complementary information. Taking this into account, one way to go ahead would be to combine Gabor and LBP methods. Following this direction, Zhang et al. proposed the so called Local Gabor binary pattern histogram sequence (LGBPHS) in which multiresolution and multiorientation description of an image using Gabor filters is first computed, and then LBP histograms are computed from the Gabor features for small nonoverlapping regions and concatenated into a feature histogram. Excellent results are reported [8]. Other works have also successfully exploited the complementary of Gabor filters and LBP features by fusing the two sets of features for recognition (see references at http://www.ee.oulu.fi/research/imag/texture/lbp/bibliography/).

Summary

Feature extraction is a very important and crucial task in all biometric systems. In this context, Gabor filtering and LBP operator are powerful means for extracting complementary features and describing biometric data such as faces, irises, fingerprints, palmprints, etc. This can be attested by the large number of successful biometric applications based on these features. Gabor filters are linear bandpass filters computed for images at different orientations and scales while LBP is a nonlinear operator codifying the relationship between each center pixel and its neighborhood, thus describing a set of local primitives such as curved edges, spots, flat areas, etc. In contrast to global features which describe

the overall content and shape of the objects in the image, local features define specific information in local regions. The local feature based methods seem to be more robust against variations in pose or illumination than holistic methods.

Related Entries

► Classifier Design
► Face Descriptors
► Feature Extraction

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circuits Syst. Video Technol., Special Issue on Image- and Video-Based Biometrics **14**(1), 4–20 (2004)
2. Jain, A.K., Prabhakar, S., Hong, L., Pankanti, S.: Filterbank-based fingerprint matching. IEEE Trans. Image Process. **9**(5), 846–859 (2000)
3. Zhang, D., Kong, W., You, J., Wong, M.: Online palmprint identification. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9),1041–1050 (2003)
4. Wiskott, L., Fellous, J.M., Kuiger, N., von der Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**, 775–779 (1997)
5. Ahonen, T., Hadid, A., Pietikäinen, M.: Face description with local binary patterns: application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 2037–2041 (2006)
6. Zhao, G., Pietikäinen, M.: Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans. Pattern Anal. Mach. Intell. **29**(6), 915–928 (2007)
7. Hadid, A., Pietikäinen, M., Li, S.Z.: Learning personal specific facial dynamics for face recognition from videos. In: Analysis and Modeling of Faces and Gestures (AMFG 2007), Lecture Notes in Computer Science, vol. 4778, pp. 1–15 (2007)
8. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition. In: Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05), 786–791 (2005)
9. Gabor, D.: Theory of communication. J. Inst. Elec. Eng. (J-IEE London) **93**(26), 429–457 (1946)
10. Granlund, G.H.: In search of a general picture processing operator. Comput. Graph. Image Process. (CGIP) **2**, 155–173 (1978)
11. Daugman, J.: Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. **2**(7), 1160–1169 (1985)
12. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition **29**, 51–59 (1996)
13. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 971–987 (2002)
14. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. **15**(11), 1148–1161 (1993)
15. Lades, M., Vorbrüggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R.P., Konen, W.: Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput. **42**, 300–311 (1993)

Local Surface Patch

A "local surface patch" (LSP) is defined as the region consisting of a feature point P and its neighbors N. The LSP representation includes feature point P, its surface type, centroid of the patch, and a histogram of shape index values vs. dot product of the surface normal at point P and its neighbors. A local surface patch is shown in Fig. 4. The neighbors satisfy the following conditions,

$$N = \{\text{pixels } N, \parallel N - P \parallel \leq \text{Є}_1\} \quad \text{and a cos}(n_p \bullet n_n < A), \tag{1}$$

where $\bullet$ denotes the dot product between the surface normal vectors n_p and n_n at point P and N and acos denotes the inverse cosine function. The two parameters Є1 and A (Є1 = 5.8 mm, $A = 0.5$) are important since they determine the descriptiveness of the local surface patch representation. A local surface patch is not computed at every pixel in a range image, but only at selected feature points. The feature points are defined as the local minimum and the maximum of shape indexes, which can be calculated from principal curvatures.

► Ear Biometrics, 3D

Localization

► Human Detection and Tracking

Localization Inaccuracy

► Face Misalignment Problem

Logical Access Control, Client-Based

Most platforms and peripherals that come with embedded fingerprint readers include software to access the local PC and applications. These applications may include biometric-based access to the PC, pre-boot authentication, full disk encryption, Windows logon, and a general password manager application to facilitate the use of biometric for other applications and websites. Such a suite of applications protects the specific PC on which it is deployed and makes personal access to data more secure, convenient, and fun. Companies such as Dell, Lenovo, Microsoft, and Hewlett-Packard ship platforms and peripherals are pre-loaded with such capability. However, these are end-user utilities with the scope of use only on the local PC. As a result, they may be challenging and costly to manage if deployed widely in an enterprise since each user will need to setup, enroll her biometric, and configure the appropriate policy, all by herself. Usually the user is given the option to use the biometric system as a cool individual convenience, rather than enforced by an enterprise-wide authentication policy.

► Access Control, Logical

Logical Access Control, Client-Server-Based

The client-server-based logical access control solutions typically limit the flexibility given to the end-user and instead focus on the needs of the organization and the system administrator to deploy, enroll users' biometric credentials into the enterprise directory, and centrally configure enterprise-wide policies. An enterprise-wide policy, however, drives stronger requirements for the reliability, security, and interoperability of the biometric authentication. If it is a business policy that everyone in the organization must use the biometric system for authentication, the reliability of the biometric system must be higher than a client-side-only solution where the user can opt-in to use the biometric system just for convenience. A server-based logical access control solution generally needs to be interoperable with data coming from many different biometric readers since not every platform in the organization will use the same model of the biometric reader. Interoperability can be accomplished at either the enrollment template level or the biometric image level. Lastly, since a server-based solution typically stores biometric credentials in a central database, the security model of the whole chain from the reader to the server must be considered to protect against hackers and maintain user privacy. However, unlike government deployments that store the user's actual biometric image(s) for archival purposes, a biometric solution used for enterprise authentication typically stores only the biometric enrollment templates.

► Access Control, Logical

Logico-Linear Operator

An operation in signal processing that bridges the gulf between linear operations, such as filtering, and the logical calculus of Boolean operators such as AND, OR, and XOR. In doing so, a logico-linear operator serves as a kind of *signal-to-symbol converter*. The input to the operator is a continuous signal such as a sound waveform or an image, upon which a linear operation is performed such as computing some derivative, or convolving with some filter. The output of the linear operation is converted into a logical state by, for example, comparing to a threshold, noting its sign, quantizing its phase or its modulus if complex, or some more abstract binary classification. The resulting Boolean quantity can be used as a logical operand for purposes such as detecting similarity and differences (XOR), motion between image frames (XOR), region growing (OR), masking (AND) of some data by other data,

descriptions of complexity or of graph structure, vetoing (NAND), machine learning, and so forth. Iris encoding and recognition is performed through the use of logico-linear operators.

► Iris Encoding and Recognition using Gabor Wavelets

Logon, Password Management

► Access Control, Logical

Luminance

In photometry, luminance is a measure of the density of illumination describing the amount of light that passes through or is emitted from a particular area, and falls within a given solid angle. In the context of color perception, luminance indicates the perceived brightness (or lightness) of a given color. In color spaces that separate the luminance in a separate channel (such as the Y channel in the YUV color space), the luminance channel of an image is equivalent to a gray level version of that image.

► Gait Recognition, Motion Analysis for

M

Machine-Generated Fingerprint Classes

Fingerprints are grouped based on some similarity criteria in the feature space. Fingerprint groups are formed by machine learning from fingerprint samples in an unsupervised manner such as clustering and binning. Such fingerprint groups are called machine-generated fingerprint classes. The goal of partitioning the database into machine-generated fingerprint classes is to divide the fingerprint population into as many classes as possible while maximizing the possibility of placing the fingerprints of a same finger into a same class in a consistent and reliable way.

► Fingerprint Classification

Machine-Learning

A type of algorithm that learns from past experience to make decisions.

► Incremental Learning
► Palmprint Matching

Magnification

In optical imaging, the ratio of the dimensions of the image created by the optical system to the dimensions of the object that is imaged. The ratio can be less than one.

► Iris Device

Mahalanobis Distance

The Mahalanobis distance is based on the covariance among variables in the feature vectors which are compared. It has the advantage of utilizing group means and variances for each variable and the problems of scale and correlation inherent in the Euclidean distance are no longer an issue. When using Euclidean distance, the set of points equidistant from a given location is a sphere. The Mahalanobis distance stretches this sphere to correct the respective scales of different variables and to account for correlation among variables.

► Hand Shape
► Signature Matching

Malicious-code-free Operating System

► Tamper-proof Operating System

Manifold

Manifold is a non-empty subset M of R^N such that the neighborhood of every point $p \in M$ resembles a Euclidean space. A smooth manifold is associated with a set of homeomorphisms that map points from open subsets around every point p to points in open subsets in R^m, where ^{m} is the intrinsic dimensionality of the manifold.

► Gait Recognition, Motion Analysis for
► Manifold Learning
► Non-linear Techniques for Dimension Reduction

© 2009 Springer Science+Business Media, LLC

Manifold Embedding

Any manifold is embedded in an Euclidean space, e.g., a sphere in the 3D world is a two-dimensional manifold embedded in a three-dimensional space.

► Gait Recognition, Motion Analysis for

Manifold Learning

Philippos Mordohai[1], Gérard Medioni[2]
[1]Stevens Institute of Technology, PA, USA
[2]University of Southern California, Los Angeles, CA, USA

Definition

Manifold learning is the process of estimating the structure of a ► manifold from a set of samples, also referred to as observations or *instances*, taken from the manifold. It is a subfield of machine learning that operates in continuous domains and learns from observations that are represented as points in a Euclidean space, referred to as the ► ambient space. This type of learning, to Mitchell, is termed *instance-based* or *memory-based* learning [1]. The goal of such learning is to discover the underlying relationships between observations, on the assumption that they lie in a limited part of the space, typically a manifold, the ► intrinsic dimensionality of a manifold of which is an indication of the degrees of freedom of the underlying system.

Introduction

Manifold learning has attracted considerable attention of the machine learning community, due to a wide spectrum of applications in domains such as pattern recognition, data mining, biometrics, function approximation and visualization. If the manifolds are linear, techniques such as the Principal Component Analysis (PCA) [2] and Multi-Dimensional Scaling (MDS) [3] are very effective in discovering the subspace in which the data lie. Recently, a number of new algorithms that not only advances the state of the art, but are also capable of learning nonlinear manifolds in spaces of very high dimensionality have been reported in the literature. These include locally linear embedding (LLE) [4], Isomap [5] and the charting algorithm [6]. They aim at reducing the dimensionality of the input space in a way that preserves certain geometric or statistical properties of the data. Isomap, for instance, preserves the ► geodesic distances between all points as the manifold is "unfolded" and mapped to a space of lower dimension.

Given a set of observations, which are represented as vectors, the typical steps of processing are as follows:

- Intrinsic dimensionality estimation
- Learning structure of the manifold
- Dimensionality reduction to remove redundant dimensions, preserving the learned manifold structure.

Not all algorithms perform all steps. For instance, LLE [4] and the Laplacian eigenmaps algorithm [7] require an estimate of the dimensionality to be provided externally. The method proposed by Mordohai and Medioni [8], which is based on tensor voting, does not reduce the dimensionality of the space, but performs all operations in the original ambient space.

Recent methods used for these tasks are discussed in this essay. Dimensionality reduction is described in conjunction with manifold learning since it is often closely tied with the selected manifold learning algorithm. In addition, research on manifold learning with applications in biometrics is highlighted.

Intrinsic Dimensionality Estimation

Bruske and Sommer [9] who proposed an approach an optimally topology preserving map (OTPM) is constructed for a subset of the data. Principal Component Analysis (PCA) is then performed for each node of the OTPM on the assumption that the underlying structure of the data is locally linear. The average of the number of significant singular values at the nodes is the estimate of the intrinsic dimensionality.

Kégl [10] estimated the capacity dimension of a manifold, which is equal to the topological dimension and does not depend on the distribution of the data, using an efficient approximation based on packing numbers. The algorithm takes into account

dimensionality variations with scale and is based on a geometric property of the data. This approach differs from the PCA-related methods that employ successive projections to increasingly higher-dimensional subspaces until a certain percentage of the data is explained.

Raginsky and Lazebnik [11] described a family of dimensionality estimators based on the concept of quantization dimension. The family is parameterized by the distortion exponent and includes Kégl's method [10] when the distortion exponent tends to infinity. The authors demonstrated that small values of the distortion exponent yield estimators that are more robust to noise.

Costa and Hero [12] estimated the intrinsic dimension of the manifold and the entropy of the samples. Using geodesic-minimal-spanning trees, the method, like Isomap [5], considers global properties of the adjacency graph and thus produces a single global estimate.

The radius of the spheres is selected in such a way that they contain enough points and that the density of the data contained in them can be assumed constant. These requirements tend to underestimate of the dimensionality when it is very high.

The method proposed by Mordohai and Medioni [14] obtains estimates of local intrinsic dimensionality at the point level. Tensor voting enables the estimation of the normal subspace of the most salient manifold passing through each point. The normal subspace is estimated locally by collecting at each point votes from its neighbors. Tensor voting is a pairwise operation in which all points act as voters casting votes to their neighbors and as receivers collecting votes from their neighbors. These votes encode geometric information on the dimensionality and orientation of the local subspace of the receiver on the assumption that the voter and receiver belong to the same structure (manifold). The dimensionality of the estimated normal subspace is given by the maximum gap in the eigenvalues of a second order tensor that represents the accumulated votes at the point. The intrinsic dimensionality of the manifold at the point under consideration is computed as the dimensionality of the ambient space minus that of the normal subspace.

Manifold Learning and Dimensionality Reduction

Schölkopf et al. [15] proposed the underlying assumption of is that if the data lie on a locally linear, low-dimensional manifold, then each point can be reconstructed from its neighbors with appropriate weights. These weights should be the same in a low-dimensional space, the dimensionality of which is greater than or equal to the intrinsic dimensionality of the manifold. The LLE algorithm computes the basis of such a low-dimensional space. The dimensionality of the embedding, however, has to be given as a parameter, since it cannot always be estimated from the data. Moreover, the output is an embedding of the given data, but not a mapping from the ambient to the ▶ embedding space. The LLE is not isometric and often fails by mapping distant points close to each other.

Isomap, an extension of the MDS, developed by Tenenbaum et al. [5] uses geodesic instead of Euclidean distances and can thus be applied to nonlinear manifolds. The geodesic distances between points are approximated by graph distances. MDS is then applied to the geodesic distances to compute an embedding that preserves the property of points to be close or far away from each other. Isomap can handle points not in the original dataset, and perform interpolation, but it is limited to convex datasets.

The Laplacian eigenmaps algorithm, developed by Belkin and Niyogi [7] computes the normalized graph Laplacian of the adjacency graph of the input data, which is an approximation of the Laplace-Beltrami operator, on the manifold. It exploits locality preserving properties that were first observed in the field of clustering. The Laplacian eigenmaps algorithm can be viewed as a generalization of LLE, since the two become identical when the weights of the graph are chosen according to the criterion of the latter. As in the case of the LLE, the dimensionality of the manifold also has to be provided, the computed embeddings are not isometric and a mapping between the two spaces is not produced.

Donoho and Grimes [16] proposed the Hessian LLE (HLLE), an approach similar to the Laplacian eigenmaps. It computes the Hessian instead of the Laplacian of the graph. The authors alleged that the Hessian is better suited than the Laplacian for detecting linear patches on the manifold. The major contribution of this approach is that it proposes a global, isometric method, which, unlike the Isomap, can be applied to non-convex datasets. The need to estimate second derivatives from possibly noisy, discrete data makes the algorithm more sensitive to noise than the others approaches.

Other related research includes the charting algorithm of Brand [6], which computes a pseudo-invertible mapping of the data as well as the intrinsic dimensionality of the manifold. The dimensionality is estimated by examining the rate of growth of the number of points contained in hyper-spheres as a function of the radius. Linear patches, areas of curvature and noise can be correctly classified using the proposed measure. At a subsequent stage, a global coordinate system for the embedding is defined. This produces a mapping between the input space and the embedding space.

Weinberger and Saul [17] developed Semidefinite Embedding (SDE) which addresses manifold learning by enforcing local isometry. The lengths of the sides of triangles formed by neighboring points are preserved during the embedding. These constraints can be expressed in terms of pairwise distances and the optimal embedding can be found by semidefinite programming. The method is computationally demanding, but can reliably estimate the underlying dimensionality of the inputs by locating the largest gap between the eigenvalues of the Gram matrix of the outputs. As in the case of the authors' approach, this estimate does not require a threshold.

The method for manifold learning described [8] by Mordohai and Medioni [8] is based on inferring the geometric properties of the manifold locally via tensor voting. An estimate of the local tangent space allows one to traverse the manifold estimating geodesic distances between points and generating novel observations on the manifold. In this method it is not necessary that the manifold is differentiable, or even connected. It can process data from intersecting manifolds with different dimensionality and is very robust to outliers. Unlike most of the other approaches, the authors did not perform dimensionality reduction, but conducted all operations in the ambient space instead. If dimensionality reduction is desired for visualization or memory saving, any technique can be applied after this.

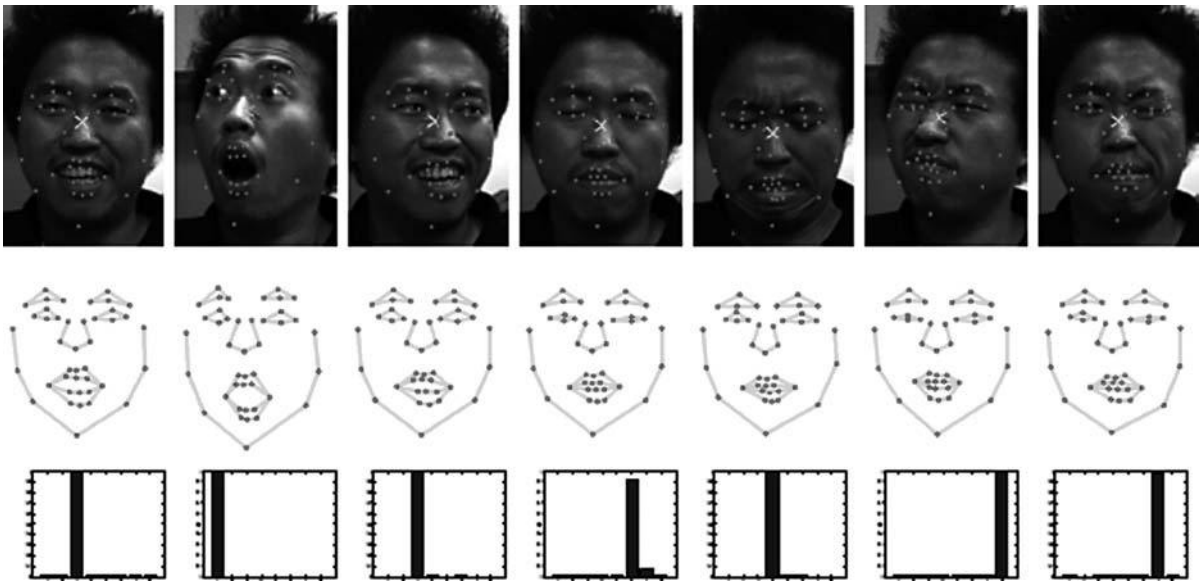
Applications

There are two main areas of application of manifold learning techniques in biometrics: estimation of the degrees of freedom of the data and visualization. Given labeled data, the degrees of freedom can be separated into those that are related to the identity of the subject and those that are due to other factors, such as pose. Visualization is enabled by reducing the dimensionality of the data to two or three to generate datasets suitable for display. This can be achieved by selecting the most relevant dimensions of the manifold and mapping them to a linear 2-D or 3-D space.

An example of both visualization and estimation of the important modes of variability of face images has been discussed by [4]. The input is a set of images of the face of a single person undergoing expression and viewpoint changes. The images are vectorized, that is the pixels of each 28*20 image are stacked to form a 560-D vector, and used as observations. LLE is able to determine the two most dominant degrees of freedom which are related to head pose and expression variations. Embedding the manifold from the 560-D ambient space to a 2-D space provides a visualization in which similar images appear close to each other. Similar experiments have been described in Tenenbaum et al. [5].

Prince and Elder [18] addressed the issue of face recognition from a manifold learning perspective by creating invariance to degrees of freedom that do not depend on identity. They labeled these degrees of freedom, namely, pose and illumination, "nuisance parameters" and were able to isolate their effects using a training dataset in which the value of the nuisance parameters is known and each individual has at least two different values of each nuisance parameter. The images are converted to 32×32 and subsequently to 1024-D vectors. Varying a nuisance parameter generates a manifold, which has little value for recognition. Therefore, once these manifolds are learned, their observations are mapped to a single point, which corresponds to the identity of the imaged person, in a new space.

Liao and Medioni [19] studied face tracking and expression inference from video sequences using tensor voting to learn manifolds that correspond to basic expressions, such as smile and surprise. During training, landmark points are tracked in the video sequence and their 3-D positions are obtained using a 3-D model of the head. Facial deformation manifolds are learned from labeled sequences of the basic expressions. A parameter that corresponds to the magnitude of the expression is estimated for each frame. During testing, the observation vector is the position of the landmarks and the goal is to jointly estimate head pose and the magnitude of each expression. This is accomplished by computing the probability that the observation was generated by each manifold. The posterior probability is inferred using a combination model of all manifolds. Some results of deformable tracking and expression inference are presented in Fig. 1.



Manifold Learning. Figure 1 *Top row*: Some frames from test video sequences [19]. *Middle row*: Visualization of the positions of the landmarks that show the estimated pose as well as the estimated deformation that corresponds to the inferred magnitude of each expression. *Bottom row*: Probability of basic expressions. Since each frame corresponds to a single expression, only one model in the mixture has a high probability.

Summary

Manifold learning techniques have attracted considerable attention in the last few years, because of their ability to untangle information in high dimensional spaces and reveal the degrees of freedom of the underlying process. This essay presents an overview of the state of the art in intrinsic dimensionality estimation and manifold learning. These algorithms can be deployed in the field of biometrics, where high dimensional data exist in large volumes, to discover and learn the dimensionality and local structure of manifolds formed by biometric measurements. Different observations on the manifold are due to variations in identity or other factors, which may be unimportant for many applications. Given training data, in which variations between samples have been labeled according to the factor that caused them, manifold learning techniques can estimate a mapping from a measurement to identity, pose, facial expression or any other variable of interest.

Related Entries

- ▶ Kernel Methods
- ▶ Machine-Learning
- ▶ Classifier Design
- ▶ Probability Distribution

References

1. Mitchell, T.: Machine Learning. McGraw-Hill, New York (1997)
2. Jolliffe, I.: Principal Component Analysis. Springer, New York (1986)
3. Cox, T., Cox, M.: Multidimensional Scaling. Chapman & Hall, London (1994)
4. Roweis, S., Saul, L.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**, 2323–2326 (2000)
5. Tenenbaum, J., de Silva, V., Langford, J.: A global geometric framework for nonlinear dimensionality reduction. Science **290**, 2319–2323 (2000)
6. Brand, M.: Charting a manifold. In: Advances in Neural Information Processing Systems 15, pp. 961–968. MIT, Cambridge, MA (2003)
7. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. **15**(6), 1373–1396 (2003)
8. Mordohai, P., Medioni, G.: Tensor Voting: A Perceptual Organization Approach to Computer Vision and Machine Learning. Morgan & Claypool, San Rafael, CA (2006)
9. Bruske, J., Sommer, G.: Intrinsic dimensionality estimation with optimally topology preserving maps. IEEE Trans. Pattern Analy. Mach. Intell. **20**(5), 572–575 (1998)
10. Kégl, B.: Intrinsic dimension estimation using packing numbers. In: Advances in Neural Information Processing Systems 15, pp. 681–688. MIT, Cambridge, MA (2003)

11. Raginsky, M., Lazebnik, S.: Estimation of intrinsic dimensionality using high-rate vector quantization. In: Advances in Neural Information Processing Systems 18, pp. 1105–1112. MIT, Cambridge, MA (2006)
12. Costa, J., Hero, A.: Geodesic entropic graphs for dimension and entropy estimation in manifold learning. IEEE Trans. Signal Process **52**(8), 2210–2221 (2004)
13. Levina, E., Bickel, P.: Maximum likelihood estimation of intrinsic dimension. In: Advances in Neural Information Processing Systems 17, pp. 777–784. MIT, Cambridge, MA (2005)
14. Mordohai, P., Medioni, G.: Unsupervised dimensionality estimation and manifold learning in high-dimensional spaces by tensor voting. International Joint Conference on Artificial Intelligence (2005)
15. Schölkopf, B., Smola, A., Müller, K.R.: Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput. **10**(5), 1299–1319 (1998)
16. Donoho, D., Grimes, C.: Hessian eigenmaps: new tools for nonlinear dimensionality reduction. In: Proceedings of National Academy of Science, pp. 5591–5596 (2003)
17. Weinberger, K.Q., Saul, L.K.: Unsupervised learning of image manifolds by semidefinite programming. Int. J. Comput. Vis. **70**(1), 77–90 (2006)
18. Prince, S., Elder, J.: Creating invariance to 'nuisance parameters' in face recognition. In: International Conference on Computer Vision and Pattern Recognition, II: pp. 446–453 (2005)
19. Liao, W.K., Medioni, G.: 3D face tracking and expression inference from a 2D sequence using manifold learning. In: International Conference on Computer Vision and Pattern Recognition (2008)

Manual Annotation

The manual annotation or description of an outsole involves a trained professional assigning a number of predefined pattern descriptors to the tread pattern. The palette of available descriptors is usually quite small and somewhat general or abstract in interpretation. For example, there may be descriptor terms such as wavy, linked, curved, zig zag, circular, simple geometric, and complex. This makes the annotation task quite subjective and inconsistent and hence must be complete by trained professionals.

► Footwear Recognition

Margin Classifier

► Support Vector Machine

Markerless 3D Human Motion Capture from Images

P. Fua
EPFL, IC-CVLab, Lausanne, Switzerland

Synonyms

Motion recovery 3D; Video-based motion capture

Definition

Markerless human motion capture from images entails recovering the successive 3D poses of a human body moving in front of one or more cameras, which should be achieved without additional sensors or markers to be worn by the person. The 3D poses are usually expressed in terms of the joint angles of a kinematic model including an articulated skeleton and volumetric primitives designed to approximate the body shape. They can be used to analyze, modify, and re-synthesize the motion. As no two people move in exactly the same way, they also constitute a signature that can be used for identification purposes.

Introduction

Understanding and recording human and other vertebrate motion from images is a longstanding interest. In its modern form, it goes back at least to Edward Muybridge [1] and Etienne-Jules Marey [2] in the nineteenth century. They can be considered as the precursors of human motion and animal locomotion analysis from video. Muybridge used a battery of photographic cameras while Marey designed an early "video camera" to capture motions such as the one of Fig. 1a. In addition to creating beautiful pictures, they pioneered image-based motion capture, motion analysis, and motion measurements.

Today, more than 100 years later, automating this process remains an elusive goal because humans have a complex articulated geometry overlaid with deformable tissues, skin, and loose clothing. They move constantly, and their motion is often rapid, complex, and self-occluding. Commercially available ► motion capture systems are cumbersome or

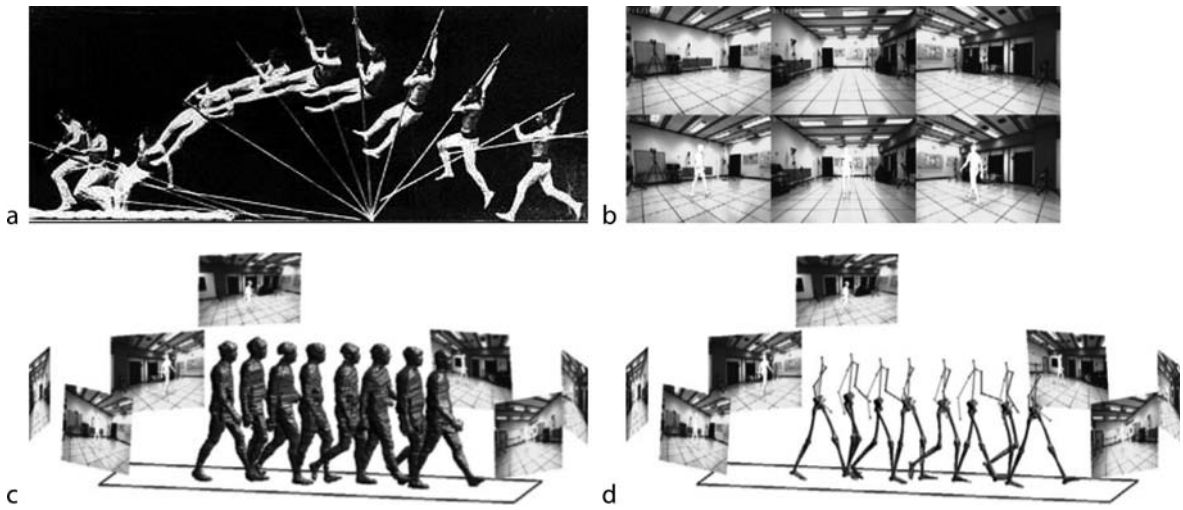
expensive or both because they rely on infra-red or magnetic sensors, lasers, or targets that must be worn by the subject. Furthermore, they usually work best in controlled environments. Markerless video-based systems have the potential to address these problems but, until recently, they have not been reliable enough to be used practically. This situation is now changing and they are fast becoming an attractive alternative.

Video-based motion capture is comparatively simpler if multiple calibrated cameras can be used simultaneously. In particular, if camera motion and background scenes are controlled, it is easy to extract the body outlines. These techniques can be very effective and commercial systems are now available. By contrast, in natural scenes with cluttered backgrounds and significant depth variation, the problem remains very challenging, especially when a single camera is used. However, it is worth addressing because solving it will result in solutions far easier to deploy and more generally applicable than the existing ones.

Success will make it possible to routinely use video-based motion capture to recognize people and characterize their motion for biometric purposes. It will also make our interaction with computers, able to perceive our gestures much more natural; allow the quantitative analysis of the movements ranging from those of athletes at sports events to those of patients whose locomotive skills are impaired; useful to capture motion sequences outside the laboratory for realistic animation and synthesis purposes; make possible the analysis of people's motion in a surveillance context; or facilitate the indexing of visual media. In short, it has many potential mass-market applications.

Methodology

This section briefly reviews the range of techniques that have been developed to overcome the difficulties inherent to 3D body motion modeling from images. This modeling is usually done by recovering the joint angles of a ▶ kinematic model that represents the subject's body, as shown in Fig. 1d. The author distinguishes between multi-camera and single-camera techniques because the former are more robust but require much more elaborate setups, which are not necessarily appropriate for biometrics applications. This section also discusses the use of ▶ pose and motion models, which have proved very effective at disambiguating difficult situations. For all the techniques introduced a few representative papers are listed. However, the author does not attempt to be exhaustive to prevent the reference list of this essay from containing several hundred entries. For a more extensive analysis, please refer [3, 4].



Markerless 3D Human Motion Capture from Images. **Figure 1** Two centuries of video-based motion capture. (**a**) Chronophotography by Marey at the end of the nineteenth century [2]. (**b**) Multi-camera setup early in the twenty-first century with background images at the top and subject's body outline overlaid in white at the bottom [3]. (**c**) Video sequence with overlaid body outlines and corresponding visual hulls [3]. (**d**) Articulated skeleton matched to the visual hulls [3].

Multi-Camera Modeling

Many methods that derive the 3D pose of a person from 3D shape sequences reconstructed from multiple views have been proposed. A popular approach is to fit a skeleton parameterized in terms of its joint angles to the visual-hull derived from body outlines [3], as illustrated by Fig. 1b–d. In a controlled environment this can be done in real-time but requires great care during the imaging process to ensure that the silhouettes can indeed be extracted reliably. An alternative is to extract stereo data using camera pairs and fitting the body model to the resulting 3D point cloud. In both cases, the process can be initialized by asking the subject to perform a sequence of known motions to estimate body proportions and calibrate the system.

Until recently, most of these approaches relied on deterministic gradient descent techniques combined with the extended Kalman filter to iteratively estimate changes in pose and motion. A common limitation of these techniques is the use of a single pose or state estimate which is updated at each time step. In practice, if the movement is too fast or if the image data can be accounted for almost as well by more than one pose, pose estimation may fail catastrophically. Monte Carlo-based tracking techniques, such as particle filtering [5], were introduced to deal with such failures by simultaneously considering multiple hypotheses in a principled way. The principal difficulty with their application to human pose estimation is the dimensionality of the state space. The number of samples or particles required increases exponentially with dimensionality. Recent work has therefore combined stochastic and gradient descent search to achieve both computational efficiency and robustness.

Techniques have also been developed to recover not only kinematic but also morphologic models, that account for body deformation during motion. These rely on machine learning approaches to perform dimensionality reduction of human shape variability and produce models [6] that can be fitted to noisy image data.

These efforts have been successful to the point where commercial systems are now becoming available. However, they usually only capture rough poses of the torso, arms, and legs while details such as hand-orientation or axial arm rotation are missing. Furthermore, the pose approximations are only dependable if the model fitted to the image data is a reasonable initial approximation of the person's body shape. The commercial systems therefore commonly assume short hair and close fitting clothing, which limits their generality.

Single Camera Modeling

Many recent approaches are trying to overcome the difficulties inherent to single-camera tracking. They can be classified as follows:

1. *Detect*. This implies recognizing postures from a single image by matching it against a database and has become increasingly popular but requires very large sets of examples to be effective. Approaches of this kind have been successfully demonstrated for pedestrian detection [7].
2. *Track*. This involves predicting the pose in a frame given the pose in the previous frame. This requires an initial pose and can easily fail if errors start accumulating in the prediction, causing divergence in the estimation process. As in the multi-camera case, this can be mitigated by introducing stochastic optimization techniques that can handle multiple competing hypotheses [5]. An effective alternative is to introduce strong dynamic motion models as priors on the search space, as will be discussed below.

Detection and tracking are complementary in many respects. They have been profitably combined to track automatically multiple people in extremely long sequences [8, 9]: Tracking takes advantage of temporal continuity and the smoothness of human motions to accumulate information through time, while detection techniques are likely to be useful for initialization of tracking and search. With suitable dynamical models, tracking has the additional advantage of providing parameter estimates that may be directly relevant for subsequent recognition tasks with applications to biometrics, sport training, physiotherapy, or clinical diagnostics.

Motion Models

Pose and motion models may be generic or activity specific. Many researchers adopt generic models that encourage smoothness while obeying kinematic joint limits. Such models are often expressed in terms of

first- or second-order Markov models. Activity-specific models more strongly constrain 3D tracking and help resolve potential ambiguities, but at the cost of having to infer the class of motion, and to learn the models.

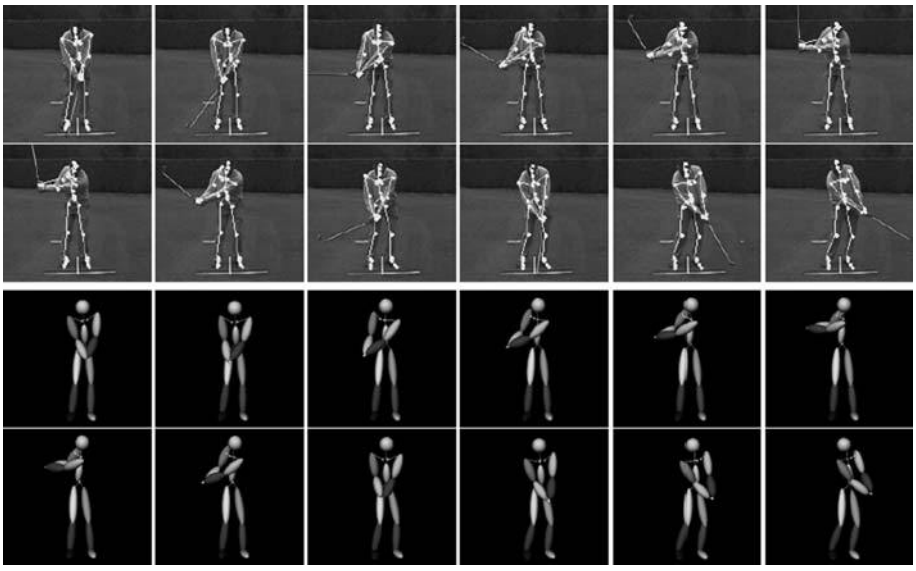
The most common approach to learning activity-specific models of motion or pose has been to use optical motion capture data from one or more people performing one or more activities, such as walking, running, or jumping. Given the high-dimensionality of the data it is natural to try embedding it in a low-dimensional space [10]. However, the highly nonlinear nature of the manifold of possible human *poses* makes it difficult. Thus, methods for nonlinear dimensionality reduction have gained in popularity. This approach is illustrated by Fig. 2 in which the motion model is expressed in terms of a Gaussian process latent variable model [11].

Instead of modeling the *pose* space, one might directly model the space of human *motions*, in which consecutive poses are concatenated into a global motion vector. Motion subspace models learned from multiple people performing the same activity have long been used in the animation community. They have also been successfully used for 3D people tracking [12, 13]. For the restricted class of cyclic motions, an automated procedure for aligning training data as a precursor to PCA was developed [12]. Similarly, a related class of subspace models for walking motions in which the temporal variations in pose is expressed in terms of sinusoidal basis functions has been proposed [14]. It has been shown that three harmonics are sufficient for reliable gender classification.

Biometric Applications

Most image-based approaches to person identification on the basis of the way they move can be classified into two broad categories: Appearance-based ones that deal directly with image statistics and model-based ones that first fit a model to the image data and then analyze the variation of its parameters.

Until now, because the model-based approaches have been so brittle, the majority of published approaches fall into the first category. Some rely on first processing each frame independently and then using a Hidden Markov Model to model the transitions from one frame to the next [15]. Others exploit the spatio-temporal statistics of the image stream [16]. Methods that rely on dense optical flow [17] or self similarity plots computed via correlation of pairs of images have also been proposed. The main drawback of these appearance-based approaches is that they are



Markerless 3D Human Motion Capture from Images. **Figure 2** Tracking of a golf swing using a single video camera. *First two rows*: The skeleton of the recovered 3D model is projected into a representative subset of images. *Bottom two rows*: Volumetric primitives of the recovered 3D model projected into the same views.

usually designed only for a specific viewpoint, usually fronto-parallel. Furthermore guaranteeing robustness against clothing and illumination changes remains difficult even though much effort has been expanded to this end, for example, by using silhouettes and binary masks rather than the image pixels gray levels themselves.

With their increasing competence, the 3D model-based approaches can be expected to eventually overcome these limitations. Already some of them have shown promise. For example, in [18], leg motion is extracted by temporal template matching using a model defined by forced coupled oscillators. Individual signatures are then derived by Fourier analysis. Another recent good example of model-based gait recognition can be found in [19]. The gait signature is extracted by using Fourier series to describe the motion of the upper leg and by applying temporal evidence gathering techniques to extract the moving model from a sequence of images. However these techniques are still 2D, which means that a near fronto-parallel view is assumed. This approach has been extended to full 3D modeling by replacing the Fourier analysis by fitting PCA-based motion models to the image data [20].

Summary

In recent years, video-based human motion capture has made very significant advances, which are driven by demands of potential mass-market applications. Multi-camera systems are beginning to reach a level of maturity that makes them of practical use but are somewhat harder to deploy and calibrate than single-camera systems. These, while still far from the robustness that would make them commercially viable, are also progressing fast. In particular, they now take advantage of sophisticated statistical learning techniques to develop effective motion models and overcome the ambiguities inherent to monocular 3D reconstruction.

Biometrics approaches are beginning to take advantage of this increasing competence to recognize people on the basis of how they move in 3D. This holds the promise of techniques that will be easy to deploy because they will only require simple and cheap sensors, such as one or more webcams, able to operate in uncontrolled environments in which the subjects can move freely.

Related Entries

- ▶ Deformable Models
- ▶ Human Detection and Tracking
- ▶ Machine-Learning
- ▶ Gait Recognition, Motion Analysis for

References

1. Muybridge, E.: Animals Locomotion. University of Pennsylvania (1887)
2. Marey, E.J.: Le mouvement. Editions Jaqueline Chambon (1994) Réédition de 1894 des éditions Masson
3. Muendermann, L., Corazza, S., Andriachhi, T.: The evolution of methods for the capture of human movement leading to markerless motion capture for biomedical applications. J. NeuroEng. Rehabil. **3** (2006)
4. Moeslund, T., Hilton, A., Krueger, V.: A survey of advances in vision-based human motion capture and analysis. Comput. Vision Image Understand **2**, 90–126 (2006)
5. Deutscher, J., Blake, A., Reid, I.: Articulated body motion capture by annealed particle filtering. In: Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC, pp. 2126–2133 (2000)
6. Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J.: Scape: shape completion and animation of people. ACM Trans. Graphics **24**, 408–416 (2005)
7. Seemann, E., Leibe, B., Schiele, B.: Multi-aspect detection of articulated objects. In: Conference on Computer Vision and Pattern Recognition, New York, NY, USA (2006)
8. Ramanan, D., Forsyth, A., Zisserman, A.: Tracking people by learning their appearance. IEEE Trans. Pattern Anal. Mach. Intell. (2007)
9. Fossati, A., Dimitrijevic, M., Lepetit, V., Fua, P.: Bridging the gap between detection and tracking for 3D monocular video-based motion capture. In: Conference on Computer Vision and Pattern Recognition, Minneapolis, MI (2007)
10. Murase, H., Sakai, R.: Moving object recognition in eigenspace representation: Gait analysis and lip reading. Pattern Recognit. Lett. **17**, 155–162 (1996)
11. Urtasun, R., Fleet, D., Hertzman, A., Fua, P.: Priors for people tracking from small training sets. In: International Conference on Computer Vision, Beijing, China (2005)
12. Ormoneit, D., Sidenbladh, H., Black, M., Hastie, T.: Learning and tracking cyclic human motion. In: Neural Information Processing Systems, pp. 894–900 (2001)
13. Sidenbladh, H., Black, M.J., Fleet, D.J.: Stochastic tracking of 3D human figures using 2D image motion. In: European Conference on Computer Vision (2000)
14. Troje, N.: Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. J. Vision **2**, 371–387 (2002)
15. He, Q., Debrunner, C.: Individual recognition from periodic activity using Hidden Markov Models. In: IEEE Workshop on Human Motion, Austin, Texas (2000)

16. Niyogi, S., Adelson, E.H.: Analyzing and recognizing walking figures in XYT. In: Conference on Computer Vision and Pattern Recognition, Seattle, WA (1994)
17. Little, J., Boyd, J.: Recognizing people by their gait: the shape of motion. Videre **1**, 1–32 (1986)
18. Yam, C.Y., Nixon, M.S., Carter, J.N.: On the relationship of human walking and running: automatic person identification by gait. In: International Conference on Pattern Recognition, Quebec, pp. 287–290 (2002)
19. Cunado, D., Nixon, M., Carter, J.: Automatic extraction and description of human gait models for recognition purposes. Comput. Vision Image Understand **90**, 1–41 (2003)
20. Urtasun, R., Fleet, D., Fua, P.: Temporal motion models for monocular and multiview 3-D human body tracking. Comput. Vision Image Understand **104**, 157–177 (2006)

Match Score Fusion

► Fusion, Score-Level

Matcher

A biometric identification system that compares the templates stored during user enrollment with those extracted from the presented biometric samples and generates a matching score. The module that generates this matching score is referred to as matcher.

► Fusion, Rank-Level

Matching

► Biometric Algorithms

Matching Score

A quantitative measure related to the similarity among a biometric trait and a user template. From a pattern classification point of view, matching scores are usually related to the likelihood of a template of being from a class. In general, given a biometric sample, the higher the matching score, the higher is the probability that it belongs to the claimed user. Matching scores are produced by the matcher module of a biometric system.

► Signature Matching

Match-On-Card

Match-on-card is a technology that enables a system to match between the template and the sampled data on a smart card. With this technology, templates will never be transmitted to any other devices or storage systems, which suppresses the risk of unauthorized template duplication. Since the matching process is executed by a relatively small CPU that is mounted on the smart card, the processing time is typically longer than other systems using common CPUs for PC.

► Finger Vein Reader
► Transportable Asset Protection

Maximum A Posteriori (MAP)

Maximum A Posteriori estimation is to estimate a stochastic variable with both prior distribution and conditional likelihood function. It can be seen as a regularization of Maximum Likelihood Estimation (MLE).

► Iris Super-Resolution

Maximum A-Posteriori Estimation

Method of parameter estimation in which a parameter is estimated using the data and a prior distribution over the parameter one wants to estimate.

► Gaussian Mixture Models

Maximum Likelihood Estimation

Method of parameter estimation in which a parameter is estimated to be that value for which the data are most likely.

► Gaussian Mixture Models

Maximum Margin Classifier

► Support Vector Machine

Maximum Permissible Exposure (MPE)

The highest exposure to which a subject may be subjected without adverse effect. Similar to but distinct from threshold limit value. Similar also to PEL, permissible exposure level, which is a legal term in some jurisdictions that defines the legally permissible exposure limit.

► Iris on the Move

Mesocephalic

Mesocephalic is the head form that is intermediate between brachycephalic and dolicocephalic forms.

► Anatomy of Face

Metatarsal Ridge

Metatarsal ridge is defined as the leading edge of the impression made by the ball area of the foot.

► Forensic Barefoot Comparisons

Microphone

► Voice Device

Microphone Arrays

Microphone arrays are composed of several microphones located at fixed relative positions from each other. They use knowledge of the microphone locations to predict the delays observed in signals coming from different directions. This allows two different possibilities: finding the position of a sound source (for instance to aim a video-conference camera at the speaker), and reducing the noise of the captured signal by enhancing the sensitivity of the microphone array in a particular direction. This is achieved by adequately combining the signals captured by the different microphones in the array.

► Voice Device

Minimal Constraint Iris Recognition

► Iris on the Move™

Minutia

Minutia are the points in the fingerprint where the finger ridges split (a bifurcation point) or terminate (an ending point).

Minutia is the basis for recognizing a fingerprint in early law enforcement applications and has gradually become the standard of fingerprint template.

► Large Scale System Design

Minutia Direction

The tangential direction of the ridge or valley at the minutia point.

► Fingerprint Features

Mislabeled Iris Data Correction

► Automatic Classification of Left/Right Iris Images

Mitochondrial DNA

Mitochondria are organelles in our cells that are associated with the production of energy. Mitochondrial DNA (mtDNA) is a circular DNA present in mitochondria and not in the cell nucleus. There are on average 100–1,000 mitochondria per cell. Each mitochondrion contains a dozen of copies of mtDNA. Therefore each cell contains 1,000–10,000 copies of mtDNA, instead of two for "normal" or nuclear DNA. mtDNA is therefore very useful for degraded samples. It is transmitted by the mother and its polymorphism is limited (the chance of finding the same sequence in two unrelated individuals is on average 1 in 1,000).

► Forensic DNA Evidence

Mixture Mode

► Gaussian Mixture Models

Model-Based Biometrics

► Biometric Sample Synthesis

Monitoring

► Surveillance

Monomodal/Multimodal Database

A monomodal database is a database which only has one biometric trait sensed. A multimodal database is a database which has more than one biometric trait from the same individual.

► Fingerprint Databases and Evaluation

Morphable Models

► Deformable Models

Mosaicing

The process of creating a composite image from overlapping component images is called mosaicing. In biometrics, mosaicing techniques are used to integrate multiple information for improving recognition performance.

► Fingerprint Templates
► Fusion, Sensor-Level

Motion Capture

Digitally recording motions by recovering the successive 3D joint angles that characterize them. Most current systems rely on specialized sensors attached to the performer's body or on imaging markers worn by the subject.

► Markerless 3D Human Motion Capture from Images

Motion Estimation

Face tracking can be looked upon as estimating the motion of the face over subsequent video frames. This can be the 2D motion on the image plane or the 3D pose of the face.

► Face Tracking

Motion Model

Motion model describes the timing and displacement of each structure such as limbs, head, and torso that makes up a particular body. The motion is dependent on the structural model. Motion displacement is normally expressed in angle or distance.

► Gait Recognition, Model-Based

Motion Recovery, 3D

► Markerless 3D Human Motion Capture from Images

Moving Light Display

A technique pioneered by Johansson [1] for use in psychological experiments to isolate a motion stimulus by acquiring images of lights placed on the body joints.

► Gait Recognition, Silhouette-Based

MS

MS is a qualitative and quantitative analytical technique which fragments compounds in a manner which is characteristic of each compound. It can be coupled with GC to produce a reliable separation and identification tool for complex mixtures.

► Odor Biometrics

Multi-Algorithm Systems

► Multibiometrics

Multi-Instance Systems

► Multibiometrics

Multi-Modal Samples

► Multibiometrics

Multi-Sample Systems

▶ Multibiometrics

Multi-Sensor Systems

▶ Multibiometrics

Multi-Unit Systems

▶ Multibiometrics

MultiBand Biometrics

▶ Multispectral and Hyperspectral Biometrics

Multibiometric Fusion, Standardization

▶ Multibiometrics and Data Fusion, Standardization

MultiBiometric Systems

The use of two or more biometrics within a biometric system to enhance its accuracy, usability, or security. The accuracy can be enhanced by leveraging the degree of orthogonality of the biometric types used, to provide an extra level of assurance the user is who he or she claims to be. The usability aspects may be enhanced by offering a type of biometric that is commensurate with the local environment: examples of this would be; using a silicon-based fingerprint system in areas of high ambient lighting; or using a facial recognition system where hands-free access is required, such as in a medical facility. The security of a system can be enhanced using a multi-biometric system not only by virtue of enhanced accuracy, but also because the use of alternate biometrics can remove the need for a non-biometric fallback system to be used in cases where one of the biometrics exhibits false rejections.

▶ Access Control, Physical

Multibiometrics

Arun Ross
Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV, USA

Synonym

Biometric fusion

Definition

Multibiometrics refers to the use of multiple sources of biometric information in order to establish the identity of an individual. Multibiometric systems combine the biometric evidence offered by multiple biometric sensors (e.g., 2D and 3D face sensors), algorithms (e.g., minutia-based and ridge-based fingerprint matchers), samples (e.g., frontal and profile face images), units (e.g., left and right irises), or traits (e.g., face and iris) to enhance the recognition accuracy of a biometric system. Information fusion can be accomplished at several different levels in a biometric system, including the sensor-level, feature-level, score-level, rank-level, or decision-level. The challenge is to design an effective fusion scheme to consolidate the multiple pieces of evidence to generate a decision about an individual's identity.

Introduction

Most biometric systems that are presently in use, typically use a single biometric trait to establish identity (i.e., they are unibiometric systems). Some of the challenges commonly encountered the by biometric systems include:

1. *Noise in sensed data.* The biometric data being presented to the system may be contaminated by noise due to imperfect acquisition conditions or subtle variations in the biometric itself.
2. *Non-universality.* The biometric system may not be able to acquire meaningful biometric data from a subset of individuals resulting in a failure-to-enroll (FTE) error.
3. *Upper bound on identification accuracy.* The matching performance of a unibiometric system cannot be indefinitely improved by tuning the feature extraction and matching modules. There is an implicit upper bound on the number of distinguishable patterns (i.e., the number of distinct biometric feature sets) that can be represented using a template.
4. *Spoof attacks.* Behavioral traits such as voice and signature are vulnerable to spoof attacks by an impostor attempting to mimic the traits corresponding to legitimately enrolled subjects. Physical traits such as fingerprints can also be spoofed by inscribing ridge-like structures on synthetic material such as gelatine and play-doh. Targeted spoof attacks can undermine the security afforded by the biometric system and, consequently, mitigate its benefits.

Some of the limitations of a unibiometric system can be addressed by designing a system that consolidates (or fuses) *multiple* sources of biometric information [1, 2]. This can be accomplished by fusing, for example, multiple traits of an individual, or multiple feature extraction and matching algorithms operating on the same biometric trait. Such systems, known as multibiometric systems [3, 4], can improve the matching accuracy of a biometric system while increasing population coverage and deterring spoof attacks. Fusion in biometrics relies on the principles in the information fusion and multiple classifier system (MCS) literature [5, 6].

Advantages of Multibiometric Systems

Besides enhancing matching accuracy, the other advantages of multibiometric systems over traditional unibiometric systems are enumerated below [3].

1. Multibiometric systems address the issue of non-universality (i.e., limited population coverage) encountered by unibiometric systems. If a subject's dry finger prevents her from successfully enrolling into a fingerprint system, then the availability of another biometric trait, say iris, can aid in the inclusion of the individual in the biometric system. A certain degree of flexibility is achieved when a user enrolls into the system using several different traits (e.g., face, voice, fingerprint, iris, hand etc.) while only a subset of these traits (e.g., face and voice) is requested during authentication based on the nature of the application under consideration and the convenience of the user.
2. Multibiometric systems can facilitate the filtering or indexing of large-scale biometric databases. For example, in a bimodal system consisting of face and fingerprint, the face feature set may be used to compute an index value for extracting a candidate list of potential identities from a large database of subjects. The fingerprint modality can then determine the final identity from this limited candidate list.
3. It becomes increasingly difficult (if not impossible) for an impostor to spoof multiple biometric traits of a legitimately enrolled individual. If each subsystem indicates the probability that a particular trait is a "spoof", then appropriate fusion schemes can be employed to determine if the user, in fact, is an impostor. Furthermore, by asking the user to present a random subset of traits at the point of acquisition, a multibiometric system facilitates a challenge-response type of mechanism, thereby ensuring that the system is interacting with a *live* user. Note that a challenge-response mechanism can be initiated in unibiometric systems also (e.g., system prompts "Please say 1-2-5-7", "Blink twice and move your eyes to the right", "Change your facial expression by smiling", etc.).
4. Multibiometric systems also effectively address the problem of noisy data. When the biometric signal

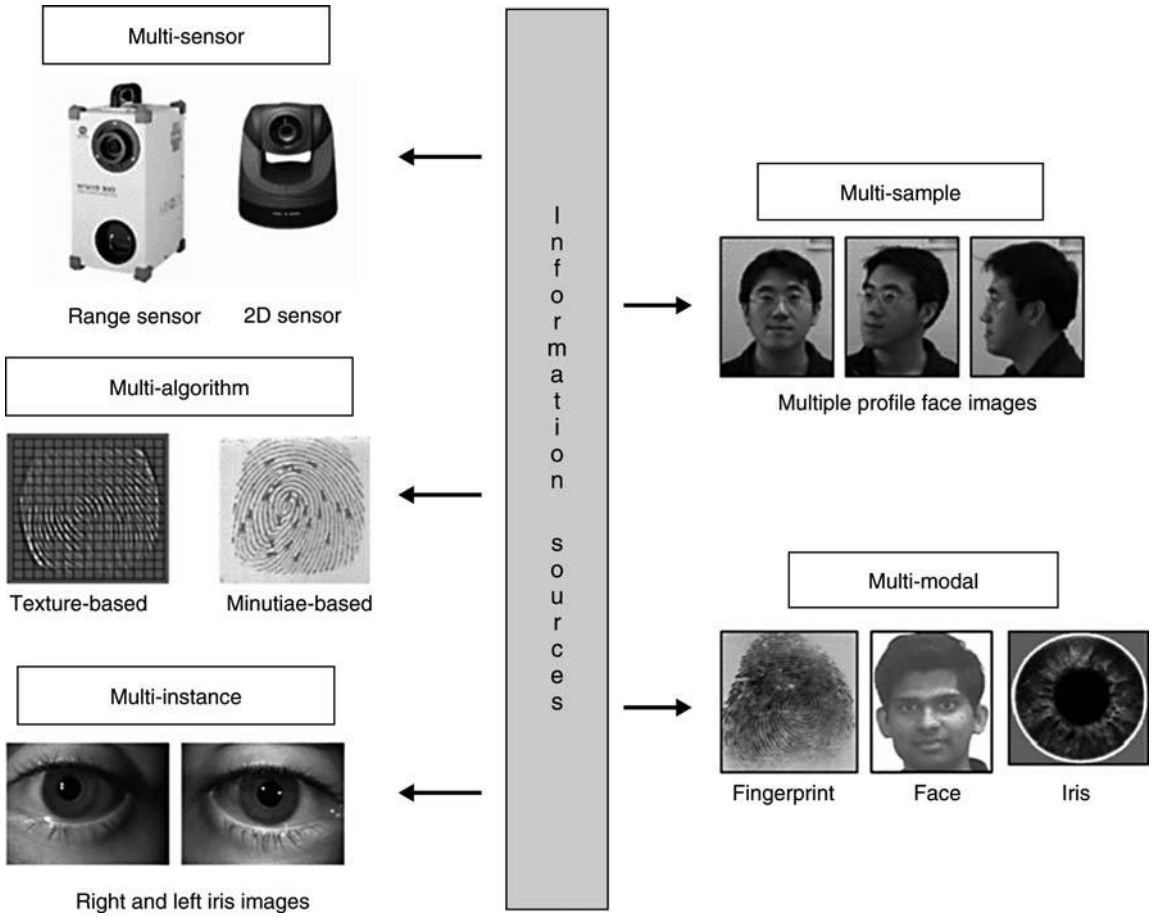
acquired from a single trait is corrupted with noise, the availability of other (less noisy) traits may aid in the reliable determination of identity. Some systems take into account the *quality* of the individual biometric signals during the fusion process. This is especially important when recognition has to take place in adverse conditions where certain biometric traits cannot be reliably extracted. For example, in the presence of ambient acoustic noise, when an individual's voice characteristics cannot be accurately measured, the facial characteristics may be used by the multibiometric system to perform authentication. Estimating the quality of the acquired data is in itself a challenging problem but, when appropriately done, can reap significant benefits in a multibiometric system.

5. These systems also help in the *continuous* monitoring or tracking of an individual in situations when a single trait is not sufficient. Consider a biometric system that uses a 2D camera to procure the face and gait information of a person walking down a crowded aisle. Depending upon the distance and pose of the subject with respect to the camera, both these characteristics may or may not be simultaneously available. Therefore, either (or both) of these traits can be used depending upon the location of the individual with respect to the acquisition system thereby permitting the continuous monitoring of the individual.
6. A multibiometric system may also be viewed as a fault tolerant system which continues to operate even when certain biometric sources become unreliable due to sensor or software malfunction, or deliberate user manipulation. The notion of fault tolerance is especially useful in large-scale authentication systems involving a large number of subjects (such as a border control application).

Taxonomy of Multibiometric Systems

A multibiometric system relies on the evidence presented by multiple sources of biometric information. Based on the nature of these sources, a multibiometric system can be classified into one of the following six categories: multi-sensor, multi-algorithm, multi-instance, multi-sample, multimodal, and hybrid (see Fig. 1).

1. *Multi-sensor systems.* Multi-sensor systems employ multiple sensors to capture a single biometric trait of an individual. For example, a face recognition system may deploy multiple 2D cameras to acquire the face image of a subject; an infrared sensor may be used in conjunction with a visible-light sensor to acquire the subsurface information of a person's face; a multispectral camera may be used to acquire images of the iris, face or finger; or an optical as well as a capacitive sensor may be used to image the fingerprint of a subject. The use of multiple sensors, in some instances, can result in the acquisition of complementary information that can enhance the recognition ability of the system. For example, based on the nature of illumination due to ambient lighting, the infrared and visible-light images of a person's face can present different levels of information resulting in enhanced matching accuracy. Similarly, the performance of a 2D face matching system can be improved by utilizing the shape information presented by 3D range images.
2. *Multi-algorithm systems.* In some cases, invoking multiple feature extraction and/or matching algorithms on the same biometric data can result in improved matching performance. Multi-algorithm systems consolidate the output of multiple feature extraction algorithms, or that of multiple matchers operating on the same feature set. These systems do not necessitate the deployment of new sensors and, hence, are cost-effective compared to other types of multibiometric systems. But on the other hand, the introduction of new feature extraction and matching modules can increase the computational complexity of these systems.
3. *Multi-instance systems.* These systems use multiple instances of the same body trait and have also been referred to as multi-unit systems in the literature. For example, the left and right index fingers, or the left and right irises of an individual, may be used to verify an individual's identity. The FBI's IAFIS (Integrated Automated Fingerprint Identification System) service combines the evidence of all ten fingers to determine a matching identity in the database. These systems can be cost-effective if a single sensor is used to acquire the multi-unit data in a sequential fashion. However, in some instances, it may be desirable to obtain the multi-unit data simultaneously



Multibiometrics. **Figure 1** Sources of information for biometric fusion.

thereby demanding the design of an effective (and possibly more expensive) acquisition device.

4. *Multi-sample systems.* A single sensor may be used to acquire multiple samples of the same biometric trait to account for the variations that can occur in the trait, or to obtain a more complete representation of the underlying trait. A face system, for example, may capture (and store) the frontal profile of a person's face along with the left and right profiles in order to account for variations in the facial pose. Similarly, a fingerprint system equipped with a small size sensor may acquire multiple dab prints of an individual's finger to obtain images of various regions of the fingerprint. A mosaicing scheme may then be used to stitch the multiple impressions and create a composite image. One of the key issues in a multi-sample system is determining the *number* of samples that have to be acquired from an individual. It is important that the procured samples represent the *variability* as well as the *typicality* of the individual's biometric data. To this end, the desired relationship between the samples has to be established beforehand to optimize the benefits of the integration strategy. For example, a face recognition system utilizing both the frontal- and side-profile images of an individual may stipulate that the side-profile image should be a three-quarter view of the face. Alternately, given a set of biometric samples, the system should be able to automatically select the "optimal" subset that would best represent the individual's variability.
5. *Multimodal systems.* Multimodal systems establish identity based on the evidence of multiple biometric traits. For example, some of the earliest multimodal biometric systems utilized face and voice features to establish the identity of an individual. Physically uncorrelated traits (e.g., fingerprint and

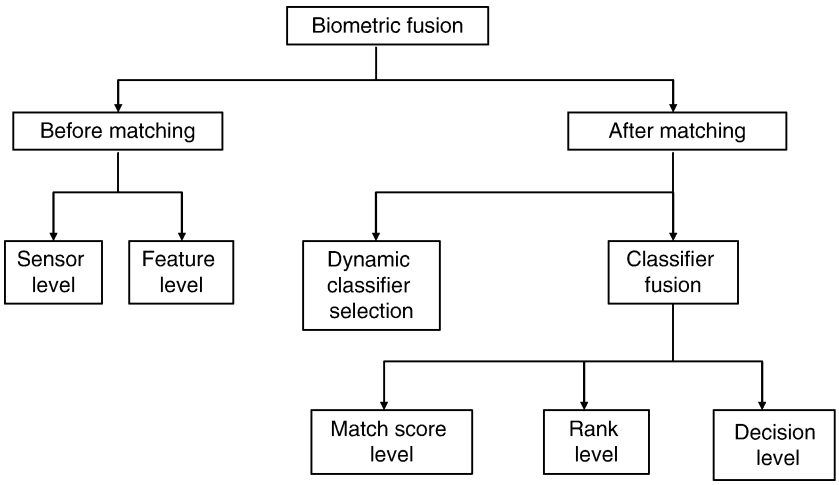
iris) are expected to result in better *improvement* in performance than the correlated traits (e.g., voice and lip movement). The cost of deploying these systems is substantially more due to the requirement of new sensors and, consequently, the development of appropriate user interfaces. The identification accuracy can be significantly improved by utilizing an increasing number of traits although the ► curse-of-dimensionality phenomenon would impose a bound on this number. The number of traits used in a specific application will also be restricted by practical considerations such as the cost of deployment, enrollment time, throughput time, expected error rate, user habituation issues, etc.

6. *Hybrid systems*. Chang et al. [7] use the term *hybrid* to describe systems that integrate a subset of the five scenarios discussed above. For example, Brunelli et al. [8] discuss an arrangement in which two speaker recognition algorithms are combined with three face recognition algorithms at the match score and rank levels via a HyperBF network. Thus, the system is multi-algorithmic as well as multimodal in its design.

Levels of Fusion

Based on the type of information available in a certain module, different levels of fusion may be defined. Sanderson and Paliwal [9] categorize the various levels of fusion into two broad categories: pre-classification or fusion *before* matching, and post-classification or fusion *after* matching (see Fig. 2). Such a categorization is necessary since the amount of information available for fusion reduces drastically once the matcher has been invoked. Pre-classification fusion schemes typically require the development of new matching techniques (since the matchers used by the individual sources may no longer be relevant) thereby introducing additional challenges. Pre-classification schemes include fusion at the sensor (or raw data) and the feature levels while post-classification schemes include fusion at the match score, rank and, decision levels.

1. *Sensor-level fusion*. The raw biometric data (e.g., a face image) acquired from an individual represents the richest source of information although it is expected to be contaminated by noise (e.g., non-uniform illumination, background clutter, etc.). Sensor-level fusion refers to the consolidation of (1) raw data obtained using multiple sensors, or (2) multiple snapshots of a biometric using a single sensor.
2. *Feature-level fusion*. In feature-level fusion, the feature sets originating from multiple biometric algorithms are consolidated into a single feature set by the application of appropriate feature normalization, transformation, and reduction schemes. The primary benefit of feature-level fusion is the detection of correlated feature values generated by different biometric algorithms and, in the process, identifying a salient set of features that can improve recognition accuracy. Eliciting this feature set typically requires



Multibiometrics. Figure 2 Fusion can be accomplished at various levels in a biometric system.

the use of dimensionality reduction methods and, therefore, feature-level fusion assumes the availability of a large number of training data. Also, the feature sets being fused are typically expected to reside in commensurate vector space in order to permit the application of a suitable matching technique upon consolidating the feature sets.

3. *Score-level fusion.* In score-level fusion the match scores output by multiple biometric matchers are combined to generate a new match score (a scalar) that can be subsequently used by the verification or identification modules for rendering an identity decision. Fusion at this level is the most commonly discussed approach in the biometric literature primarily due to the ease of accessing and processing match scores (compared to the raw biometric data or the feature set extracted from the data). Fusion methods at this level can be broadly classified into three categories: density-based schemes, transformation-based schemes, and classifier-based schemes.
4. *Rank-level fusion.* When a biometric system operates in the identification mode, the output of the system can be viewed as a ranking of the enrolled identities. In this case, the output indicates the set of possible matching identities sorted in decreasing order of confidence. The goal of rank level fusion schemes is to consolidate the ranks output by the individual biometric subsystems to derive a consensus rank for each identity. Ranks provide more insight into the decision-making process of the matcher compared to just the identity of the best match, but they reveal less information than match scores. However, unlike match scores, the rankings output by multiple biometric systems are comparable. As a result, no normalization is needed and this makes rank level fusion schemes simpler to implement compared to the score level fusion techniques.
5. *Decision-level fusion.* Many commercial off-the-shelf (COTS) biometric matchers provide access only to the final recognition decision. When such COTS matchers are used to build a multibiometric system, only decision level fusion is feasible. Methods proposed in the literature for decision level fusion include "AND" and "OR" rules, majority voting, weighted majority voting, Bayesian decision fusion, the Dempster–Shafer theory of evidence, and behavior knowledge space.

Summary

Multibiometric systems are expected to enhance the recognition accuracy of a personal authentication system by reconciling the evidence presented by multiple sources of information. Typically, early integration strategies (e.g., feature-level) are expected to result in better performance than late integration (e.g., score-level) strategies. However, it is difficult to predict the performance gain due to each of these strategies prior to invoking the fusion methodology. While the *availability* of multiple sources of biometric information (pertaining either to a single trait or to multiple traits) may present a compelling case for fusion, the ▶ correlation between the sources has to be examined before determining their suitability for fusion [10]. Combining uncorrelated or negatively correlated sources is expected to result in a better improvement in matching performance than combining positively correlated sources [11]. However, defining an appropriate diversity measure to predict fusion performance has been elusive thus far. Other topics of research in multibiometrics include (1) protecting multibiometric templates; (2) indexing multimodal databases; (3) consolidating biometric sources in highly unconstrained non-ideal environments; (4) designing dynamic fusion algorithms to address the problem of incomplete input data; and (5) predicting the matching performance of a multibiometric system.

Related Entries

▶ Fusion, Decision-Level
▶ Fusion, Feature-Level
▶ Fusion, Rank-Level
▶ Fusion, Score-Level
▶ Fusion, Sensor-Level
▶ Multiple Classifier Systems
▶ Multiple Experts

References

1. Bigun, E.S., Bigun, J., Duc, B., Fischer, S.: Expert Conciliation for Multimodal Person Authentication Systems using Bayesian Statistics. In: Proceedings of First International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 291–300. Crans-Montana, Switzerland (1997)
2. Hong, L., Jain, A.K., Pankanti, S.: Can Multibiometrics Improve Performance? In: Proceedings of IEEE Workshop on Automatic

Identification Advanced Technologies (AutoID), pp. 59–64. New Jersey, USA (1999)
3. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. 1st ed. Springer, New York, USA (2006)
4. Jain, A.K., Ross, A.: Multibiometric Systems. Communications of the ACM, Special Issue on Multimodal Interfaces **47**(1), 34–40 (2004)
5. Kuncheva, L.I.: Combining Pattern Classifiers - Methods and Algorithms. Wiley (2004)
6. Ho, T.K.: Multiple Classifier Combination: Lessons and Next Steps. In: H. Bunke, A. Kandel (eds.) Hybrid Methods in Pattern Recognition, *Machine Perception and Artificial Intelligence*, vol. 47, pp. 171–198. World Scientific (2002)
7. Chang, K.I., Bowyer, K.W., Flynn, P.J.: An Evaluation of Multimodal 2D+3D Face Biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(4), 619–624 (2005)
8. Brunelli, R., Falavigna, D.: Person Identification Using Multiple Cues. IEEE Transactions on Pattern Analysis and Machine Intelligence **17**(10), 955–966 (1995)
9. Sanderson, C., Paliwal, K.K.: Information Fusion and Person Verification Using Speech and Face Information. Research Paper IDIAP-RR 02-33, IDIAP (2002)
10. Poh, N., Bengio, S.: How Do Correlation and Variance of Base-Experts Affect Fusion in Biometric Authentication Tasks?. IEEE Transactions on Signal Processing **53**(11), 4384–4396 (2005)
11. Kuncheva, L.I., Whitaker, C.J., Shipp, C.A., Duin, R.P.W.: Is Independence Good for Combining Classifiers? In: Proceedings of International Conference on Pattern Recognition (ICPR), vol. 2, pp. 168–171. Barcelona, Spain (2000)

Multibiometrics and Data Fusion, Standardization

Jung Soh[1], Farzin Deravi[2], Alessandro Triglia[3], Alex Bazin[4]
[1]University of Calgary, Calgary, AB, Canada
[2]University of Kent, Canterbury, Kent, UK
[3]OSS Nokalva, Inc., Somerset, NJ, USA
[4]Fujitsu Services, London, UK

Synonyms

Biometric fusion standardization; Multibiometric fusion standardization

Definition

Multibiometrics is the automated recognition of individuals based on their biological or behavioral characteristics and involves the use of biometric fusion. Some applications of biometrics require a level of technical performance that is difficult to obtain with a single biometric measure. Preventing illegitimate multiple applications by the same individual for national identity cards and checking security for air travel are examples of such applications. In addition, provision is needed for people who are unable to give a reliable biometric sample for some biometric modalities. Use of multiple biometric measurements from substantially independent biometric sensors, algorithms, or modalities typically gives improved technical performance, increases system flexibility and reduces security risks. This includes an improved level of performance where not all biometric measurements are available such that decisions can be made from any number of biometric measurements within an overall policy on accept/reject thresholds. At the current level of understanding, combining results from different biometric sources at the matching score level typically requires knowledge of both genuine and impostor distributions of such scores. Such distributions are highly application-dependent and generally unknown in any real system. Research on the methods not requiring previous knowledge of the score distributions is continuing and research on fusion at both the image and feature levels is still progressing. Preliminary work on ISO/IEC international standardization of multibiometrics has culminated in a Technical Report, while in the United States substantial progress has been made on standards to support multibiometrics.

Overview of Multibiometric Systems

In general, the use of the terms ▶ multimodal or ▶ multibiometric indicates the presence and use of more than one ▶ biometric modality, sensor, instance, and/or algorithm in some form of combined use for making a specific biometric identification or verification decision. The methods of combining multiple samples, matching scores, or matching decisions can be very simple or mathematically complex.

Multimodal biometrics were first proposed, implemented and tested in the 1970s. Combining measures was seen as a necessary future requirement for biometric systems. It was widely thought that combining multiple measures could increase either security by decreasing the false acceptance rate or user convenience by decreasing the false rejection rate. These

early systems, however, did not seem to advance into practical applications.

The use of fusion and related methods has been a key tool in the successful implementation of large-scale automated fingerprint identification systems (AFISs), starting in the 1980s. Until recently, multiple modalities have not been used in AFIS; however, many methods of fusion have been successfully implemented using fingerprints alone. Some of the ways that fusion has been implemented in AFISs include:

1. Image (i.e., sample) fusion in creating a single "rolled" image from a series of plain impressions on a livescan device.
2. Template fusion in the use of multiple feature extraction algorithms on each fingerprint image.
3. Multiinstance fusion in the use of fingerprints from all ten fingers.
4. Multipresentation fusion in the use of rolled and slap (plain) fingerprints.
5. Algorithm fusion for the purpose of efficiency (cost, computational complexity, and throughput rate); generally matchers are used as a series of filters in order of increasing computational complexity. These are generally implemented as a mix of decision- and score-level fusion.
6. Algorithm fusion for the purpose of accuracy (decreasing false accept rate and/or false reject rate, lessening sensitivity to poor-quality data); matchers are used in parallel, with fusion of resulting scores.

The use of fusion has made AFISs possible, because of fusion's increase in both accuracy and efficiency. To further understand the distinction among the multibiometric categories, Table 1 illustrates the basic distinctions among categories of multibiometric implementation. The key aspect of the category that makes it multi-"something" is shown in boldface.

Multimodal biometric systems take input from single or multiple sensors that capture two or more biometric characteristics of different modalities. For example, a single system combining face and iris information for biometric recognition would be considered a "multimodal" system regardless of whether face and iris images were captured by different imaging devices or the same device. It is not required that the various measures be mathematically combined in any way. For example, a system with fingerprint and voice recognition would be considered "multimodal" even if the "OR" rule was being applied, allowing users to be verified using either of the modalities.

Multialgorithmic biometric systems receive a single sample from a single sensor and process that sample with two or more algorithms. This technique could be applied to any modality. Maximum benefit (theoretically) would be derived from algorithms that are based on distinctly different and independent principles (such algorithms may be called "orthogonal").

Multiinstance biometric systems use one (or possibly multiple) sensor(s) to capture samples of two or more different instances of the same biometric characteristic. For example, systems capturing images from multiple fingers are considered to be multiinstance rather than multimodal. However, systems capturing, for example, sequential frames of facial or iris images are considered to be *multipresentation* rather than multiinstance.

Multisensorial biometric systems sample the same instance of a ▶ biometric characteristic with two or more distinctly different sensors. Processing of the multiple samples can be done with one algorithm, or

Multibiometrics and Data Fusion, Standardization. **Table 1** Multibiometric categories illustrated by the simplest case of using 2 of something

Category	Modality	Algorithm	Biometric characteristic (e.g., body part)	Sensor
Multimodal	**2 (always)**	2 (always)	2 (always)	2 (usually)[a]
Multialgorithmic	1 (always)	**2 (always)**	1 (always)	1 (always)
Multiinstance	1 (always)	1 (always)	**2 instances of 1 characteristic (always)**	1 (usually)[b]
Multisensorial	1 (always)	1 (usually)[c]	1 (always, and same instance)	**2 (always)**
Multipresentation	1	1	1	1

[a]Exception: a multimodal system with a single sensor used to capture two different modalities (e.g., a high resolution image used to extract face and iris or face and skin texture)

[b]Exception may be the use of two individual sensors each capturing one instance (e.g., possibly a two-finger fingerprint sensor)

[c]It is possible that two samples from separate sensors could be processed by separate "feature extraction" algorithms, and then through a common comparison algorithm, making this "1.5 algorithms," or two completely different algorithms

some combination of multiple algorithms. For example, a face recognition application could use both a visible light camera and an infrared camera coupled with a specific frequency (or several frequencies) of infrared illumination.

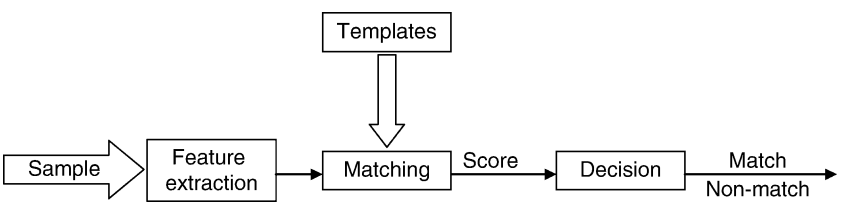
For a specific application in an operational environment, there are numerous system design considerations, and trade-offs that must be made, among factors such as improved performance (e.g., identification or verification accuracy, system speed and throughput, robustness, and resource requirements), acceptability, ease of circumvention, ease of use, operational cost, environmental flexibility, and population flexibility [1]. Especially for a large-scale human identification system, there are additional system design considerations such as operation and maintenance, reliability, system acquisition cost, life cycle cost, and planned system response to identified susceptible means of attack, all of which will affect the overall deployability of the system [1].

Levels of Combination

As a basis for the definition of levels of combination in multibiometric systems, Fig. 1 shows a single-biometric process. A biometric *sample* captured by a biometric sensor (e.g., a fingerprint image) is fed into the *feature extraction* module. Using signal processing methods, the feature extraction module converts a sample into features (e.g., fingerprint minutiae), which form a representation apt for matching. Usually, multiple features are collected into a feature vector. The *matching* module takes the feature vector as input and compares it to a stored *template* (a type of biometric reference, as defined in [2]). The result is a match *score*, which is used by the *decision* module to decide (e.g., by applying a threshold) whether the presented sample matches with the stored template. The outcome of this decision is a binary match or non-match. Generalizing the above process to a multibiometric one, there are several levels at which fusion can take place: (1) decision level; (2) match score level; (3) feature level; and (4) sample level.

Decision-level fusion takes place only after the results of matching from all biometric components are available. The decision module outputs match or non-match as a binary decision value. If a biometric system consists of a small number of biometric components, assigning logical values to match outcomes allows fusion rules to be formulated as logical functions. For two decision-level outputs, two most commonly used logical functions are logical AND and OR. For many decision-level outputs, various voting schemes can be used as fusion rules, the most common of which is majority voting. The logical AND and OR functions can be considered as voting schemes.

In *score-level* fusion, each system provides matching scores indicating the proximity of the feature vector with the template vector. These scores can then be combined to improve the matching performance. The match score output by a matcher contains the richest information about the input biometric sample in the absence of feature-level or sensor-level information. Furthermore, it is relatively easy to access and combine the scores generated by several different matchers. Consequently, integration of information at the match score level is the most common approach in multimodal biometric systems. From a theoretical point of view, biometric processes can be combined reliably to give a guaranteed improvement in matching performance. Any number of suitably characterized biometric processes can have their matching scores combined in such a way that the multibiometric combination is guaranteed (on average) to be no worse than the best of the individual biometric devices. The key is to

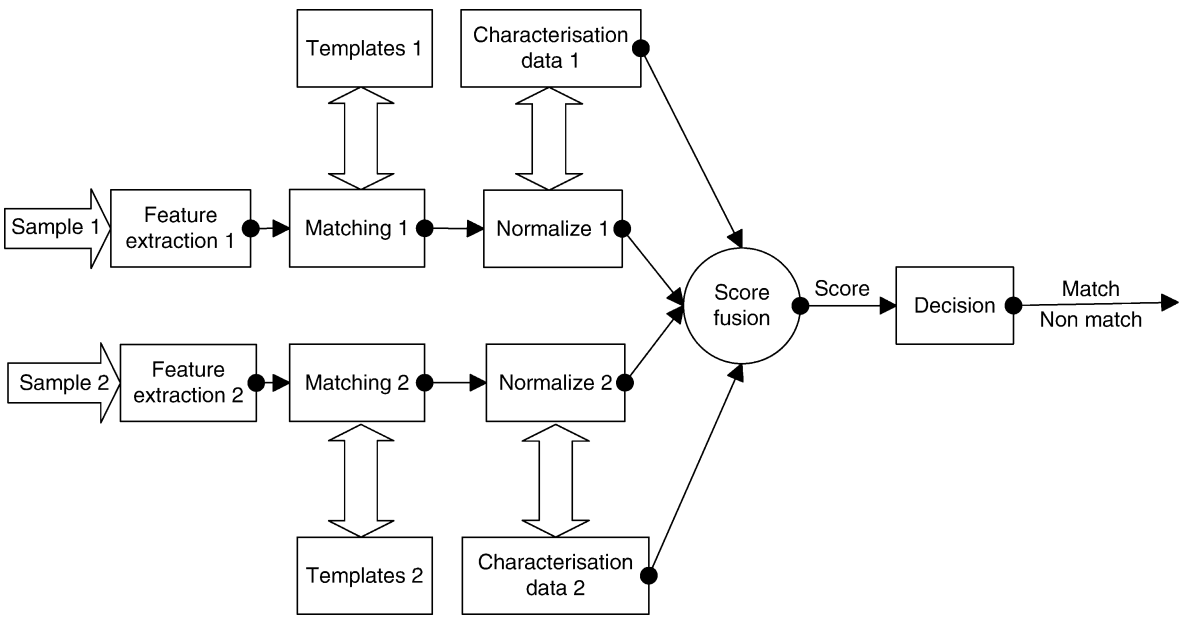


Multibiometrics and Data Fusion, Standardization. **Figure 1** Generic single-biometric process. (Reproduced from Figure 2, Section 5.1 of ISO/IEC TR 24722:2007 Information technology – Biometrics – Multimodal and other multibiometric fusion).

identify correctly the method which will combine these matching scores reliably and maximize the improvement in matching performance. The mechanism (for this sort of good combination of scores within a multibiometric system) must follow at least two guidelines: (1) each biometric process must produce a score and make it available to the multibiometric combiner; and (2) in advance of operational use, each biometric process must make available to the multibiometric combiner, its technical performance (such as score distributions) in the appropriate form (and with sufficient accuracy of characterization). Both verification (1:1) and identification (1:N) systems can support fusion at the match score level.

In the context of verification, there are two distinct approaches to formulate a score-level fusion problem: (1) classification; and (2) combination [3]. In the classification approach, a feature vector is constructed using the matching scores output by the individual matchers; this feature vector is then classified into one of two classes: "Accept" (genuine user) or "Reject" (impostor). Generally, the classifier used for this purpose (e.g., decision tree, neural network, support vector machine, *k*-nearest neighbor, random forest, etc.) is capable of learning the decision boundary, given some training data, irrespective of how the feature vector is generated [4]. Hence, the output scores of the different modalities can be non-homogeneous (distance or similarity metric, different numerical ranges, etc.) and no processing is required prior to presenting them to the classifier. In the combination approach (see Fig. 2), the individual matching scores are combined to generate a single scalar score, which is then used to make the final decision [5]. To ensure a meaningful combination of the scores from the different modalities, if necessary, the scores may be first transformed to a common domain prior to combining them. This is known as score normalization.

Score normalization methods attempt to map the scores of each biometric process to a common domain. Some reasons why scores need to be normalized prior to fusion include [3]: (1) the matching scores at the output of the individual matchers may not be homogeneous. For example, one matcher may output a distance (dissimilarity) measure while another may output a proximity (similarity) measure; (2) the outputs of the individual matchers need not be on the same numerical scale (range); and (3) the matching scores at the output of the matchers may follow different statistical distributions. Some approaches are based on the Neyman-Pearson lemma [6], with simplifying assumptions. For example, mapping scores to likelihood ratios allows them to be combined by multiplying under an



Multibiometrics and Data Fusion, Standardization. **Figure 2** A framework for score-level fusion using combination approach. (Reproduced from Figure 5, Section 5.3.2 of ISO/IEC TR 24722:2007 Information technology – Biometrics – Multimodal and other multibiometric fusion).

independence assumption. Other approaches may be based on modifying other statistical measures of the match score distributions. The parameters used for normalization can be determined using a fixed training set or adaptively based on the current feature vector.

In *feature-level* fusion, biometric information is fused after feature extraction but before matching. The simplest form is to integrate the feature vectors (or sets if there is no implicit correspondence) of component biometrics and to apply feature classification methods to the combined feature vector. Where features from contributing multibiometrics are not independent, good feature-level combination should, in some circumstances, allow dependencies to be more fully exploited than by solely using score-level combination. This should give better overall performance. However, fusion at this level is difficult to achieve in practice because of the following reasons: (1) the feature vectors of multiple modalities may be incompatible (e.g., minutiae set of fingerprints and Eigen-coefficients of face); (2) the relationship between the feature spaces of different biometric systems may not be known; (3) concatenating two feature vectors may result in a feature vector with very large dimensionality leading to the "curse of dimensionality"; and (4) a significantly more complex matcher might be required in order to operate on the concatenated feature vector [7].

Notwithstanding these challenges, fusion at the feature level has been attempted in several contexts. Chang et al. [8] demonstrate feature-level fusion of face and ear modalities showing significant improvements in performance. Kumar et al. [9] integrate the palm-print and hand geometry features of an individual in order to enhance matching performance. In their experiments, fusion at the match score level was observed to be superior to fusion at the feature level. However, Ross and Govindarajan [2] combine the hand and face modalities of a user (multibiometrics) as well as the R, G, B channels of the face image (multisensorial) of a user at the feature level and demonstrate that a feature selection scheme may be necessary to improve matching performance at this level.

State of International Standardization of Multibiometrics

At the time of this writing (late 2008), no standards on multibiometrics have been developed within ISO/IEC. However, the ISO/IEC subcommittee on biometrics (ISO/IEC JTC 1/SC 37) had instead produced a Technical Report on multibiometrics [10] (hereafter referred to as the "Technical Report"), which contains descriptions and analyses of current practices on multibiometric fusion and provides useful information for future development of international standards in this area.

According to the Technical Report, there are many ways of combining multibiometric processing and performing ▶ biometric fusion, not all of which can be made part of a biometric fusion standard. It is likely that future biometric fusion standardization activity within ISO/IEC will be of five types:

1. *Record formats:* The definition and standardization of data to be exchanged between processes and stored on various media. The biometric data interchange formats specified in SC 37/WG 3 standards are examples of this type of standard. Another example is the Fusion Information Format national standard developed in the US [11]
2. *Interfaces:* Definition of standard APIs for processes, the record formats used by the processes, and the initialization procedure of the processes in a system. The BioAPI standard [12] developed in SC 37/WG 2 is an example of this type of standard, which might have to be modified in order to support multibiometrics and fusion data. The US version of the BioAPI standard [13] has been amended [14] to support biometric fusion
3. *Application profile:* A standard containing a list of references to provisions of one or more other standards, which are specified as optional in those standards but are made mandatory by this standard in order to facilitate interoperability in a particular set of use cases. The SC 37/WG 4 project on ILO (International Labor Organization) Seafarer ID profile [15] is an example of this type of standard
4. *Conformance testing:* A description of the criteria and test data that allows for the assurance that systems have complied with the standards. These types of standards are under development in SC 37 for the biometric record formats
5. *Performance testing:* Online testing of biometric systems is complicated by the implied existence of multiple and sequential sensors. A testing protocol that develops procedures for doing this should be established

The use of multibiometric systems has been considered for two major and differing use cases. The first is high-security biometric use where the combination of biometrics provides a stronger assurance of impostor rejection for a relatively small, trained population. The second is in the context of large-scale ID systems, such as travel document systems, where the multibiometric combination may provide for the reduction of rejection rates and easier system usage for a very large, untrained population. In the context of the large-scale ID systems, there can be many solution providers providing components to the overall system. For example, the creator of the electronic biometric document may not be the same vendor that creates the physical document, and neither may be the vendor that performs the biometric test(s) (verification or identification) during the document's usage. This situation can clearly benefit from a biometric fusion standard when the document contains multiple biometrics.

In the context of biometric fusion, one can propose the following multibiometric system interoperability requirements:

1. Standard multibiometric systems may be required to be designed and certified (or evaluated) based on common performance requirements. These performance requirements should be independent of the biometric modalities in use. This includes performance measures such as failure to enroll, failure to acquire, false rejection rate, false acceptance rate, system throughput, and the resistance to active impostor attacks
2. Standard multibiometric systems may be required to be designed so that a single biometric subsystem can be separately upgraded. All biometric device characteristics change over time as research and development improves accuracy and lowers cost. The development of each biometric system however, proceeds on its own timeline. Therefore only if separate upgrading is possible it will be convenient to upgrade a multibiometric system in the field
3. A standard multibiometric system may be required to be able to accept historical information for a given user, such as scores and processing times. With this information, the system can be optimized in both security and throughput to take advantage of the type of biometric modality that is favored by the particular user
4. Standard multibiometric systems may be required to be compatible with existing standard-based systems that use a single biometric characteristic. To achieve support of the system requirements, existing biometric technical interfaces, such as the BioAPI standard, may need to be revised to provide support for fusion while allowing the use of independently developed BioAPI Biometric Service Providers (BSPs) each implementing a single biometric modality

State of Standardization of Multibiometrics in the United States

In the United States a significant amount of work had been done, at the time of this writing, to produce the first two national standards on multibiometrics. The two standards that have been developed so far are:

1. An amendment [14] to the US version of the BioAPI standard [13] that adds support for biometric fusion (published)
2. A data format carrying information in support of score-level fusion [11] (published)

The changes made to BioAPI 1.1 to support fusion were inspired by the Technical Report. The Technical Report describes a multibiometric process as a combination of processes, and describes the inputs and the outputs of each process. This general model is now reflected in BioAPI 1.2.

In BioAPI 1.2, the concept of Biometric Information Record (BIR) was generalized, and now includes the following types of biometric data (sometimes called "processed levels"):

1. Raw biometric data (e.g., a raw image produced for audit purposes during a capture operation)
2. Intermediate biometric data (e.g., an image ready to be passed as input to a feature extraction process)
3. Processed data (e.g., features)
4. A matching score (as produced by some matching algorithm)
5. A matching decision (a match/non-match output)
6. Personal data (any non-biometric data specific to the subject that can be passed as input to a biometric operation)

Correspondingly, the BioAPI operations that handle biometric data (either accepting one or more BIRs as

input or producing a BIR as output) have been extended in BioAPI 1.2 to support the new types of BIRs. For example, an important evolution was the addition of a function similar to VerifyMatch but producing a BIR of type "score" instead of the traditional "achieved FAR" (a simple integer value) and "result" (a simple boolean value). In BioAPI 1.2, the new Fuse function can take as input two or more "score" BIRs to produce either a new "score" BIR or a "decision" BIR (score-level fusion); it can also take as input two or more "decision" BIRs to produce a new "decision" BIR (decision-level fusion) or two or more "processed" BIRs to produce a new "processed" BIR (feature-level fusion). The Fuse function can be implemented either by a regular BSP that also performs capture and/or matching, or by a specialized "fusion BSP." The new functions also accept as input any number of additional BIRs of any type (including "personal") carrying information that can be used by the BSP to increase the quality of the output of the function.

The Fusion Information Format standard [11] specifies a container of information about the statistical distribution of similarity scores for a particular biometric technology (including a matcher). The standard supports the representation of both genuine and impostor distributions, and provides four different ways of representing those distributions: location and scale parameters, empirical cumulative distribution function, B-spline function fit of the empirical cumulative distribution function (CDF), and interpolant of the CDF. This information is intended to be provided as input to a software component that performs score-level fusion within a multibiometric system. When such a software component is invoked to fuse the scores produced by two or more comparison subsystems, the knowledge of the score distributions of each comparison subsystems will allow it to correctly process the corresponding input scores.

Summary

The building blocks for a biometric fusion standard would be mainly of two types: data records and processes. The three key factors driving the implementation of the fusion algorithms will be: interoperability, performance, and industry consensus. Sample-, feature-, score-, and decision-level fusion have been identified from the preliminary work on biometric fusion standardization. However, supporting sample- or feature-level fusion will be a challenge. The nature of feature-level fusion requires the definition and creation of a feature specific to a particular biometric characteristic and capture/extraction system, as well as a matching algorithm for the fused feature. Requiring vendors to support feature-level fusion across many biometric modalities may not be practical, given the current level of industry consensus reachable in today's marketplace. On the contrary, decision-level fusion is rather simple mathematically, so a fusion standard might not seem to be required for this level of fusion. Yet the initialization, security specification, and using multiple biometric decisions make it an inherently complicated process, and there is significant benefit to be gained by including decision-level fusion in standards. Nonetheless, the first beneficiaries of the fusion standardization activity are most likely score-level fusion systems. For this reason, the ISO/IEC Technical Report provides a wealth of descriptions of score-level fusion, and the current US multibiometrics standardization work focuses mainly on supporting score-level fusion.

Acknowledgment

The terms and definitions taken from ISO/IEC TR 24722:2007 Information technology – Biometrics – Multimodal and other multibiometric fusion, sections 4.1, 5.1, 5.3.1, 5.3.2, 5.3.3, 5.4, 7.2, 7.3 and the introduction, are reproduced with permission of the International Organization for Standardization, ISO. This standard can be obtained from any ISO member and from the Web site of ISO Central Secretariat at the following address: www.iso.org. Copyright remains with ISO.

Related Entries

- ▶ Fusion, Decision-Level
- ▶ Fusion, Feature-Level
- ▶ Fusion, Rank-Level
- ▶ Fusion, Sensor-Level
- ▶ Multi-Algorithm Systems
- ▶ MultiBiometrics
- ▶ Multi-Instance Systems
- ▶ Multi-Modal Systems

- ▶ Multi-Sample Systems
- ▶ Multi-Sensor Systems
- ▶ Multi-Unit Systems
- ▶ Multiple Classifier Systems
- ▶ Multiple Experts
- ▶ Score-Level Fusion

References

1. Korves, H., Nadel, L., Ulery, B., Masi, D.: Multibiometric fusion: From research to operations. Mitretek Sigma, Summer 2005 (2005)
2. ISO/IEC JTC 1/SC 37 Standing Document 2 – Harmonized Biometric Vocabulary, ISO/IEC JTC 1/SC 37 N2263 (2007)
3. Jain, A.K., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. Pattern Recognit. **38**(12), 2270–2285 (2005)
4. Wang, Y., Tan, T., Jain, A.K.: Combining face and iris biometrics for identity verification. In: Proceedings of the Fourth International Conference on Audio- and Video-Based Biometric Person Authentication, Guildford, UK, pp. 805–813 (2003)
5. Kittler, J., Hatef, M., Duin, R.P., Matas, J.G.: On combining classifiers. IEEE Trans Pattern Anal. Mach. Intell. **20**(3), 226–239 (1998)
6. Neyman, J., Pearson, E.S.: On the problem of the most efficient tests of statistical hypotheses. Philos. Trans. R. Soc. Lond A. **231**, 289–337 (1933)
7. Ross, A., Govindarajan, R.: Feature level fusion using hand and face biometrics. In: Proceedings of the SPIE Conference on Biometric Technology for Human Identification, Orlando, USA, pp. 196–204 (2005)
8. Chang, K., Bowyer, K.W., Sarkar, S., Victor, B.: Comparison and combination of ear and face images in appearance-based biometrics. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1160–1165 (2003)
9. Kumar, A., Wong, D.C.M., Shen, H.C., Jain, A.K.: Personal verification using palmprint and hand geometry biometric. In: Proceedings of the Fourth International Conference on Audio- and Video-Based Biometric Person Authentication, Guildford, UK, pp. 668–678 (2003)
10. ISO/IEC TR 24722:2007, Information technology – Biometrics – Multimodal and other multibiometric fusion (2007)
11. ANSI INCITS 439, Information Technology – Fusion Information Format for Data Interchange (2008)
12. ISO/IEC 19784-1:2006, Information technology – Biometric Application Programming Interface – Part 1: BioAPI Specification (2006)
13. ANSI INCITS 358–2002, Information technology – BioAPI Specification (2002)
14. ANSI INCITS 358-2002/AM1-2007, Information technology – BioAPI Specification – Amendment 1: Support for Biometric Fusion (2007)
15. ISO/IEC 24713–3, Information technology – Biometric profiles for interoperability and data interchange – Part 3: Biometric based verification and identification of Seafarers (2008)

Multifactor

Multifactor authentication/identification solutions consist of a combination (integrated or loosely linked) of different categories of authentication and identification technologies. A multifactor solution could thus be composed of a fingerprint recognition system tied to a proximity card reader and a PIN-based keypad system.

- ▶ Fraud Reduction, Overview

Multimodal

- ▶ Multibiometrics

Multimodal Fusion

- ▶ Multiple Experts

Multimodal Jump Kits

A compact, durable, and mobile kit that encases multiple biometric testing devices. An example kit is a briefcase that contains digital fingerprints, voice and iris prints, and photographs.

- ▶ Iris Acquisition Device

Multiple Classifier Fusion

- ▶ Fusion, Decision-Level

Multiple Classifier Systems

Fabio Roli
Department of Electrical and Electronic Engineering, University of Cagliari, Piazza d'Armi, Cagliari, Italy

Synonyms

Classifier combination; Ensemble learning; Multiple classifiers; Multiple expert systems

Definition

The rationale behind the growing interest in multiple classifier systems is the acknowledgment that the classical approach to design a pattern recognition system that focuses on finding the best individual classifier has some serious drawbacks. The most common type of multiple classifier system (MCS) includes an ensemble of classifiers and a function for parallel combination of classifier outputs. However, a great number of methods for creating and combining multiple classifiers have been proposed in the last 15 years. Although reported results showed the good performances achievable by combining multiple classifiers, so far a designer of pattern classification systems should regard the MCS approach as an additional tool to be used when building a single classifier with the required performance is very difficult, or does not allow exploiting the complementary discriminatory information that other classifiers may encapsulate.

Motivations of Multiple Classifiers

The traditional approach to classifier design is based on the "evaluation and selection" method. Performances of a set of different classification algorithms are assessed against a data set, name validation set, and the best classifier is selected. This approach works well when a large and representative data set is available, so that the estimated performances allow selecting the best classifier for future data collected during the operation of the classifier machine. However, in many real cases where only small training sets are available, estimated performances can substantially differ from the ones that classifiers will exhibit during their operation. This is the well-known phenomenon of the generalization error, which can makes impossible the selection of the best individual classifier, or cause the selection of a classifier with a poor performance. In the worst case, the worst classifier in the considered ensemble could exhibit the best apparent accuracy when assessed against a small validation set.

A first motivation for the use of multiple classifiers comes from the intuition that instead of selecting a single classifier, a safer option would be to use them all and "average" their outputs [1, 2]. This combined classifier might not be better than the individual best classifier, but the combination should reduce the risk of selecting a classifier with poor performance. Experimental evidences and theoretical results support this motivation. It has been proved that averaging the outputs of multiple classifiers do eliminates the risk of selecting the worst classifier, and can provide a performance better than the one of the best classifier under particular conditions [3].

Dietterich suggested further reasons for the use of a multiple classifier system (MCS) [2]. Some classifiers, such as neural networks, are trained with algorithms that may lead to different solutions, that is, different classification accuracies, depending on the initial learning conditions. Combining multiple classifiers obtained with different initial learning conditions (e.g., different initial weights for a neural net), reduces the risk of selecting a classifier associated to a poor solution of the learning algorithm (a so called "local optimum"). The use of MCS can simplify the problem of choosing adequate values for some relevant parameters of the classification algorithm (e.g., the number of hidden neurons in a neural net). Multiple versions of the same classifier with different values of the parameters can be combined. In some practical cases with small training sets, training and combining an ensemble of simple classifiers (e.g., linear classifiers) to achieve a certain high accuracy can be easier than training directly a complex classifier [4]. Finally, for some applications, such as multi-modal biometrics, the use of multiple classifiers is naturally motivated by application requirements.

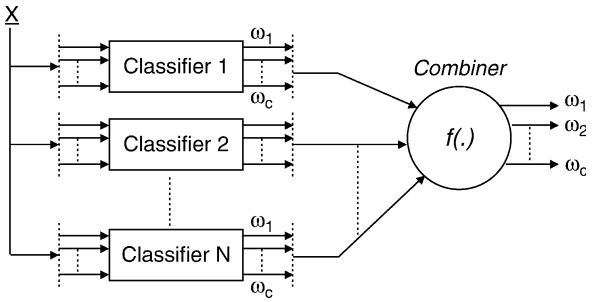
It is worth noting that the above motivations neither guarantee that the combination of multiple classifiers always performs better than the best individual classifier in the ensemble, nor an improvement on the ensemble's average performance for the general case. Such guarantees can be given only under particular conditions that

classifiers and the combination function have to satisfy [3]. However, reported experimental results and theoretical works developed for particular combination functions show the good performances achievable by combining multiple classifiers. So far a designer of pattern classification systems should regard the MCS approach as an additional tool to be used when building a single classifier with the required performance is very difficult, or does not allow exploiting the complementary discriminatory information that other classifiers may encapsulate.

Design of Multiple Classifier Systems

The most common type of MCS, widely used and investigated, includes an ensemble of classifiers, named ► "base" classifiers, and a function for parallel combination of classifier outputs (Fig. 1). The base classifiers are often algorithms of the same type (e.g., decision trees or neural networks), and statistical classifiers are the most common choice. The use of hybrid ensembles containing different types of algorithms has been investigated much less, as well as ensembles of structural classifiers have not attracted much attention, though they could be important for some real applications.

The design of a MCS involves two main phases: the design of the classifier ensemble and the design of the combination function [4]. Although this formulation of the design problem leads one to think that effective design should address both the phases, most of the design methods described in the literature focused



Multiple Classifier Systems. **Figure 1** Standard architecture of a multiple classifier system for a classification task with *c* classes. The multiple classifier system is made up by an ensemble of *N* classifiers and a function for parallel combination of classifier outputs.

on only one. Two main design approaches have been proposed, that Ho called "coverage optimization" and "decision optimization" methods [5]. Coverage optimization refers to methods that assume a fixed, usually simple, decision combination function and aim to generate a set of mutually complementary classifiers that can be combined to achieve optimal accuracy. Several techniques have been proposed for creating a set of mutually complementary classifiers. The main approaches are outlined in the following. However, decision optimization methods assume a given set of carefully designed classifiers and aim to select and optimize the combination function. These methods fit well with those applications where a set of classifiers developed separately is already available (e.g., a face and a fingerprint classifier in biometric applications) and one is interested in combining them optimally. A large set of combination functions of increasing complexity is available to the designer to perform the selection and optimization, ranging from simple voting rules through "trainable" combination functions. The three main types of combination functions are briefly explained below. There are few methods for building a MCS that cannot be classified as either a decision optimization or a coverage optimization method. For example, MCS based on the mixture of experts model trains the classifiers and the combination function simultaneously, so implementing a sort of joint optimization [6]. In some works dealing with real-life applications, hybrid design methods that first generate a classifier ensemble and then select and optimize the combination function have been also proposed.

It is easy to see that the classifiers in a MCS should be as accurate as possible and should not make coincident errors. Although this sounds an intuitive and simple concept, it revealed a complex issue that was addressed in the literature under the name of classifier "diversity" [1, 7]. The type of diversity required for maximizing MCS performance obviously depends on the combination function used. As an example, coincident errors can be tolerated if the majority-voting rule is used, but the majority of classifier decisions should be always correct. Many diversity measures have been proposed, and some of them have been also used to design effective MCSs [4]. The main approaches for creating multiple classifiers, which are outlined in the following, aim to induce classifier diversity. However, so far it appeared to be difficult to

define diversity measures that are related well to the MCS performance, so that such performance can be predicted by measuring the diversity of the classifiers.

Although the parallel architecture was the most used and investigated (Fig. 1), other types of architectures are also possible. For example, serial architectures where classifiers are applied in succession, with each classifier working on inputs, which previous classifiers were not able to recognize with a sufficient confidence. Such architectures could be very important for real applications that demand for a trade-off between accuracy and computational complexity. However, non-parallel architectures have been relatively neglected.

Although some design methods proved to be very effective and some works investigated the comparative advantages of different methods [4, 5], clear guidelines are not yet available for choosing the best design method for the classification task at hand. The designer of MCS has a toolbox containing quite a lot of instruments for generating and combining classifiers. She/he may design a myriad of different MCSs by coupling different techniques for creating classifier ensembles with different combination functions. However, for the general case, the best MCS can only be determined by performance evaluation. Optimal design is possible only under particular assumptions on the classifiers and the combination function [3].

Creating Classifier Ensembles

Several techniques have been proposed for creating a set of complementary classifiers. All these techniques try to induce classifier diversity, namely, to create classifiers that make errors on different patterns, so that they can be combined effectively. In the following, the main approaches to classifier ensemble generation are outlined.

Using Different Base Classifiers

A simple way for generating multiple classifiers is using base classifiers of different types. For example, classifiers based on different models (e.g., neural networks and decision trees) or using different input information. This simple technique may work well for applications where complementary information sources are available (e.g., multi-sensor applications) or distinct representations of patterns are possible (e.g., minutiae-based and texture-based representations in fingerprint classification).

Injecting Randomness

Random variation of some parameters of the learning or classification algorithm can be used to create multiple classifiers. The classifier ensemble is created using multiple versions of a certain base classifier obtained by random variation of some parameters. For example, training a neural network several times with different random values of the initial weights allows generating a network ensemble.

Manipulating Training Data

These techniques generate multiple classifiers by training a base classifier with different data sets. To this end, the most straightforward method is the use of disjoint training sets obtained by splitting the original training set (this technique is called sampling without replacement). A very popular technique based on training data manipulation is Bagging (Bootstrap AGGregatING) [8]. Bagging creates an ensemble made of N classifiers trained on N bootstrap replicates of the original training set. A bootstrap replicate consists of a set of m patterns drawn randomly with replacement from the original training set of m patterns. The classifiers are usually combined by majority voting rule, or by averaging their outputs. Another popular technique based on training data manipulation is Boosting [9]. This method incrementally builds a classifier ensemble. The classifier that joins the ensemble at step k is forced to learn patterns that previous classifiers have misclassified. In other words, while Bagging samples each training pattern with equal probability, Boosting focuses on those training patterns that are most often misclassified. Essentially, a set of weights is maintained over the training set and adaptive resampling is performed, such that the weights are increased for those patterns that are misclassified. It is worth noting that Boosting is a complete design method, where the combination function is also specified, and it is not only a technique for generating a classifier ensemble.

Manipulating Input Features

Manual or automatic feature selection can be used for generating multiple classifiers using different feature sets as inputs. Ho proposed a successful technique of this type, called Random Subspace Method [10]. Feature space is randomly sampled, such that complementary classifiers are obtained by training them with different feature sets.

Manipulating Output Labels

A multiclass problem can be subdivided into a set of subproblems (e.g., two-class problems), and a classifier can be associated to each subproblem, thereby generating an ensemble. The standard problem subdivision is the so-called "one-per-class" decomposition, where each classifier in the ensemble is associated to one of the classes and it is aimed to discriminate such class from the others. Dietterich and Bakiri proposed a general method, called ECOC (error correcting output codes) method, for generating multiple classifiers by decomposing a multiclass task into subtasks [11].

Combining Multiple Classifiers

There are two main strategies in combining classifiers: fusion and selection [1]. The most of the combination functions follow one of these basic strategies, with the majority of combination methods using the fusion strategy. There are few combiners that use hybrid strategies, where fusion and selection are merged. In classifier fusion, each classifier contributes to the final decision for each input pattern. In classifier selection, each classifier is supposed to have a specific domain of competence (e.g., a region in the feature space) and is responsible for the classification of patterns in this domain. There are combination functions lying between these two main strategies. For example, the mixture of experts model uses a combination strategy that, for each input pattern, selects and fuses a subset of available classifiers [6]. Some combination functions, such as the majority-voting rule, are called "fixed" combiners because they do not need training (i.e., they do not need estimation of parameters from a training set). Other combination functions need additional training and they are called "trainable" combiners. For example, the weighted majority-voting rule needs to estimate weights that are used to give different importance to the classifiers in the vote. Trainable combiners can obviously outperform the fixed ones, supposed that a large enough and independent validation set for training them in an effective way is available [4]. It should be noted that each classifier in the ensemble is often biased on the training data, so that the combiner should not be trained on such data. An additional data set that is independent from the training set used for the individual classifiers should be used. In general, the complexity of the combiner should be adapted to the size of the data set available. Complex trainable combiners, that need to estimate a lot of parameters, should be used only when large data sets are available [1, 12].

Fusion of Multiple Classifiers

The combination functions following the fusion strategy can be classified on the basis of the type of outputs of classifiers forming the ensemble. Xu et al. distinguish between three types of classifier outputs [13]: (1) Abstract-level output: Each classifier outputs a unique class label for each input pattern; (2) Rank-level output: Each classifier outputs a list of ranked class labels for each input pattern. The class labels are ranked in order of plausibility of being the correct class label; (3) Measurement-level output: Each classifier outputs a vector of continuous-valued measures that represent estimates of class posterior probabilities or class-related confidence values that represent the support for the possible classification hypotheses. On the basis of this classification, the following three classes of fusion rules can be defined.

Abstract-Level Fusion Rules

Among the fusion rules that use only class labels to combine classifier outputs, the most often used rule is the majority vote that assigns the input pattern to the majority class, that is, the pattern is assigned to the most frequent class in the classifier outputs. A natural variant of the majority vote, namely, the plurality vote, is also used. The trainable version of the majority vote rule is the weighted majority vote that uses weights that are usually related to the classifier performances. Among the trainable fusion rules of this type, a popular rule is the Behavior Knowledge Space (BKS) method [1]. In the BKS method, every possible combination of abstract-level classifiers outputs is regarded as a cell in a look-up table. The BKS table is designed by a training set. Each cell contains the samples of the training set characterized by a particular value of class labels. Training samples in each cell are subdivided per class, and the most representative class label (the "majority" class) is selected for each cell. For each unknown test pattern, the classification is performed according to the class label of the BKS cell indexed by the classifier outputs. The BKS method requires very large and representative data sets to work well.

Rank-Level Fusion Rules

The most commonly used rule of this type is the Borda Count method. The Borda count method combines

the lists of ranked class labels provided by classifiers, and it classifies an input pattern by its overall class rank, that is computed summing the rank values that classifiers assigned to the pattern for each class. The class with the maximum overall rank is the winner. Rank-level fusion rules are suitable for problems with many classes, where the correct class may appear often near the top of the list, although not always at the top.

Measurement-Level Fusion Rules

Examples of fixed rules that combine continuous classifier outputs are: the simple mean (average), the maximum, the minimum, the median, and the product of classifier outputs [1]. Linear combiners (i.e., the average and its trainable version, the weighted average) are used in popular ensemble learning algorithms such as Bagging [8], the Random Subspace Method [10], and AdaBoost [1, 9], and represent the baseline and first choice combiner in many applications. Continuous classifier outputs can be also regarded as a new feature space (an intermediate feature space [1]). Another classifier that takes classifier outputs as input and outputs a class label can do the combination. However, this approach usually demands very large data sets that allow training effectively this additional classifier. The Decision Templates method is an interesting example of a trainable rule for combining continuous classifier outputs [1]. The idea behind decision templates combiner is to store the most typical classifier outputs (called decision template) for each class, and then compare it with the classifier outputs obtained for the input pattern (called decision profile of the input pattern) using some similarity measure.

Selection of Multiple Classifiers

In classifier selection, the role of the combiner is selecting the classifier (or the subset of classifiers) to be used for classifying the input pattern, under the assumption that different classifiers (or subsets of classifiers) have different domains of competence. Dynamic classifier selection rules have been proposed that estimate the accuracy of each classifier in a local region surrounding the pattern to be classified, and select the classifier that exhibits the maximum accuracy [14, 15]. As dynamic selection may be too computationally demanding and require large data sets for estimating the local classifier accuracy, some static selection rules have also been proposed where the regions of competence of each classifier are estimated before the operational phase of the MCS [1]. Classifier selection has not attracted as much attention as classifier fusion, probably due to the practical difficulty of identifying the domains of competence of classifiers that make possible an effective selection.

Related Entries

- ▶ Ensemble Learning
- ▶ Fusion, Decision-Level
- ▶ Fusion, Rank-Level
- ▶ Fusion, Score-Level
- ▶ Multi-Algorithm Systems
- ▶ Multiple Experts

References

1. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms, Wiley, NY (2004)
2. Dietterich, T.G.: Ensemble methods in machine learning, Multiple Classifier Systems, Springer-Verlag, LNCS, **1857**, 1–15 (2000)
3. Fumera, G., Roli, F.: A Theoretical and experimental analysis of linear combiners for multiple classifier systems. IEEE Trans. Pattern Anal. Mach. Intell. **27**(6), 942–956 (2005)
4. Roli F., Giacinto, G.: Design of multiple classifier systems. In: Bunke, H., Kandel, A. (eds.) Hybrid Methods in Pattern Recognition, World Scientific Publishing (2002)
5. Ho, T.K.: Complexity of classification problems and comparative advantages of combined classifiers, Springer-Verlag, LNCS, **1857,** 97–106 (2000)
6. Jacobs, R., Jordan, M., Nowlan, S., Hinton, G.: Adaptive mixtures of local experts. Neural Comput. **3**, 79–87 (1991)
7. Kuncheva, L.I., Whitaker, C.J.: Measures of diversity in classifier ensembles. Mach. Learn. **51**, 181–207 (2003)
8. Breiman, L.: Bagging predictors. Mach. Learn. **24**, 123–140 (1996)
9. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. **55**(1), 119–139 (1997)
10. Ho, T.K.: The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Intell. **20**, 832–844 (1998)
11. Dietterich, T.G., Bakiri, G.: Solving multiclass learning problems via error-correcting output codes. J. Artif. Intell. Res. **2**, 263–286 (1995)
12. Roli, F., Raudys, S., Marcialis, G.L.: An experimental comparison of fixed and trained fusion rules for crisp classifiers outputs. In: Proceedings of the third International Workshop on Multiple Classifier Systems (MCS 2002), Cagliari, Italy, June 2002, LNCS **2364**, 232–241 (2002)

13. Xu, L., Krzyzak, A., Suen, C.Y.: Methods for combining multiple classifiers and their applications to handwriting recognition. IEEE Trans. Syst. Man Cybern. **22**(3), 418–435 May/June (1992)
14. Woods, K., Kegelmeyer, W.P., Bowyer, K.: Combination of multiple classifiers using local accuracy estimates. IEEE Trans. Pattern Anal. Mach. Intell. **19**(4), 405–410 (1997)
15. Giacinto, G., Roli, F.: Dynamic Classifier Selection Based on Multiple Classifier Behavior. Pattern Recognit. **34**(9), 179–181 (2001)

Multiple Classifiers

► Multiple Classifier Systems

Multiple Expert Systems

► Multiple Classifier Systems

Multiple Experts

Josef Bigun
Halmstad University, IDE, Halmstad, Sweden

Synonyms

Decision fusion; Feature fusion; Fusion; Multimodal fusion; Score fusion

Definition

A *Biometric expert*, or an expert in biometric recognition context, refers to a method that expresses an opinion on the likelihood of an identity by analyzing a signal that it is specialized on, e.g. a fingerprint expert using minutiae, a lip-motion expert using statistics of optical-flow. Accordingly, there can be several experts associated with the same sensor data, each analyzing the data in a different way. Alternatively, they can be specialized on different sensor data. Multiple experts can address the issue of how expert opinions should be represented and reconciled to a single opinion on the authenticity of a ► client identity.

Introduction

In biometric signal analysis, the fusion of multiple experts can in practice be achieved as ► feature fusion or *score fusion*. In addition to these, one can also discern *data fusion*, e.g. stereo images of a face, and *decision fusion*, e.g. the decisions of several experts wherein each expresses either of the crisp opinions "client" or "► impostor," in the taxonomy of fusion [1]. However, one can see data fusion and decision fusion as adding novel experts and as a special case of score fusion, respectively. On the other hand, feature fusion is often achieved as concatenation of feature vectors, which is in turn modeled by an expert suitable for the processing demands of the set of the novel vectors. For this reason we only discuss score fusion in this article. The initial frameworks for fusion have been simplistic in that no knowledge on the skills of the experts is used by the ► supervisor. Later efforts to reconcile different expert opinions in a multiple experts biometric system have been described from a probabilistic opinion modeling [2] and a pattern discrimination [3], view points, respectively. From both perspectives, it can be concluded that the weighted average is a good way of reconciling different authenticity scores of individual experts to a single opinion, under reasonable conditions. As the weights reflect the skills of the experts, some sort of training is needed to estimate them. Belonging to probabilistic modeling school, respective discriminant analysis school, Bayesian modeling [4, 5], and support vector machines [6–8] have been utilized to fuse expert opinions. An important issue for a fusion method is, however, whether or not it has mechanisms to discern the general skill of an expert from the quality of the current data. We summarize the basic principles to exemplify typical fusion approaches as follows.

Simple Fusion

This type of fusion applies a rule to input opinions delivered by the experts. The rule is not obtained by

training on expert opinions, but are decided by the designer of the supervisor. Assuming that the supervisor receives all expert inputs in parallel, common simple fusions include,

max	Maximum of the scores,	$M_j = \max(x_{1,j}, x_{2,j}, \ldots, x_{m,j})$
min	Minimum of the scores,	$M_j = \min(x_{1,j}, x_{2,j}, \ldots, x_{m,j})$
sum	Arithmetic mean of the scores,	$M_j = \frac{1}{m}\sum_{i=1}^{m} x_{i,j}$
median	Median of the scores,	$M_j = x_{\left(\frac{m+1}{2}\right)j}$
Product	Geometric mean of the scores,	$M_j = (\prod_{i=1}^{m} x_{i,j})^{\frac{1}{m}}$

where M_j is the score output by the supervisor at the instant of operation j, when m expert opinions, expressed as real numbers $x_{i,j}, i: 1\ldots m$, are available to it. In addition to a parallel application of a single simple fusion to all expert opinions, one can apply several simple fusion rules serially (one after the other) if some expert opinions are delayed before they are processed by the supervisor(s).

Probabilistic Fusion

Experts can express opinions in various ways. The simplest is to give a strict decision on a claim of an identity, "1" (client) or "0" (impostor). A more common way is to give a graded opinion, usually a real number in [0, 1]. However, it turns out that machine experts can benefit from a more complex representation of an opinion, an array of real variables. This is not surprising to human experience because, a human opinion is seldom so simple or lacks variability that it can be described by what a single variable can afford. A richer representation of an opinion is therefore the use of the distribution of a score rather than a score. Bayes theory is the natural choice in this case because it is about how to update knowledge represented as distribution (prior) when new knowledge (likelihood) becomes available.

Before describing a particular way of constructing a Bayesian supervisor let us illustrate the basic mechanism of Bayesian updating. Let two stochastic variables X_1, X_2 represent these errors of two different measurement systems measuring the same physical quantity. We assume that these errors are independent and are distributed normally as $N(0, \sigma_1^2)$, $N(0, \sigma_2^2)$, respectively. Then their weighted average

$$M = q_1 X_1 + q_2 X_2, \qquad \text{where} \quad q_1 + q_2 = 1 \tag{1}$$

is also normally distributed with $N(0, {q_1}^2\sigma_1^2 + {q_2}^2\sigma_2^2)$. Given the variances σ_1^2, σ_2^2, if the weights q_1, q_2 are chosen inversely proportional to the respective variances, the variance of the new variable M (the weighted mean) will be smallest provided that

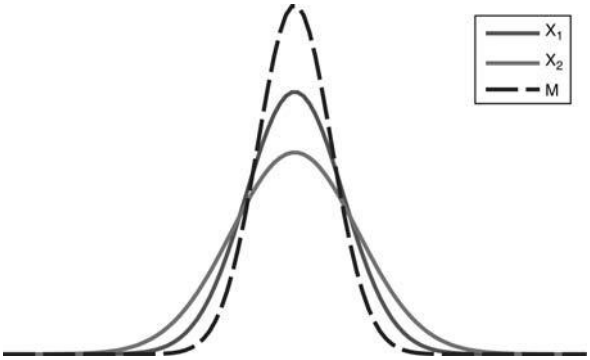
$$q_1 = \frac{\frac{1}{\sigma_1^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}, \qquad q_2 = \frac{\frac{1}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \tag{2}$$

where inverse-proportionality constants (the denominators) ensure $q_1 + q_2 = 1$. Notice that the composite variable M is normally distributed always if the X_1, X_2 are independent but the variance is smallest only for a particular choice, (seen earlier) yielding

$$\mathrm{var}(M) = q_1^2\sigma_1^2 + q_2^2\sigma_2^2 = \frac{\frac{1}{\sigma_1^4}}{\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^2}\sigma_1^2 + \frac{\frac{1}{\sigma_2^4}}{\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^2}\sigma_2^2 = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \le \min(\sigma_1^2, \sigma_2^2) \tag{3}$$

The fact that the composite variance never exceeds the smallest of the component variances, and that it converges to the smallest of the two when either becomes large, i.e. one distribution approaches the noninformative distribution $N(0, \infty)$, can be exploited to improve the precision of the aggregated measurements, Fig. 1.

Appropriate weighting is the main mechanism on how knowledge as represented by distributions can be utilized to improve biometric decision making. Bayes theory comes handy at this point because it offers the powerful Bayes theorem to estimate the weights for the aggregation of the distributions, incrementally, or at one-go as new knowledge becomes available. We follow [4] to



Multiple Experts. Figure 1 The component distributions with $X_1 \sim N(0,1)$, $X_2 \sim N(0, 1.3^2)$ and the composite distribution $M \sim N(0, 1/(1.3^{-2}+1))$, (1).

exemplify the Bayesian approach. Let the following list describe the variables representing the signals made available by a multiexpert biometric system specialized in making decisions. Next we will discuss the errors of the experts on client and impostor data separately.

- i: Index of the experts. $i \in 1 \ldots m$,
- j: Index of shots (one or more per candidate), $j \in 1 \ldots n$, n_T. It is equivalent to time since an expert has one shot per evaluation time (period). The time n is the last instant in the training whereas n_T is the test time when the system is in operation.
- X_{ij}: The authenticity score, i.e. the score delivered by expert i on shot j's claim of being a certain client
- Y_j The true authenticity score of shot j's claim being a certain client. This variable can take only two numerical values corresponding to "True" and "False"
- Z_{ij} The miss-identification score, that is $Z_{ij} = Y_j - X_{ij}$
- S_{ij} The variance of Z_{ij} as estimated by expert i

One can model the errors (not the scores) that a specific expert makes when it encounters clients. To this end, assume that $Y_j = 1$ and that the conditional stochastic variable Z_{ij} given its expectation value b_i is normally distributed i.e. $(Z_{ij}|\, b_i) \sim N(b_i, {\sigma_{ij}}^2)$. If Z_{ij} are independent then, according to Bayes theory, the *posterior* distribution $(b_i|\, z_{ij})$, will also be normal

$$(b_i|z_{ij}) \sim N(M_i^C, V_i^C) \tag{4}$$

with mean and variance

$$M_i^C = \frac{\sum_{j=1}^{n^C} \frac{z_{ij}}{{\sigma_{ij}}^2}}{\sum_{j=1}^{n^C} \frac{1}{{\sigma_{ij}}^2}} \quad \text{and} \quad V_i^C = \left(\sum_i^{n^C} \frac{1}{\sigma_{ij}^2} \right)^{-1} \tag{5}$$

respectively. In this updating, we see the same pattern as in the example, (1–3). Here C is a label that denotes that the applicable variables relate to clients. This distribution at hand, one can now estimate b_i as the expectation of $(b_i|\, z_{ij})$ which is ${M_i}^C$. This derivation can be seen as that we updated a noninformative prior distribution, $b_i \in N(0, \infty)$, i.e. "nothing is known about b_i" to obtain the posterior distribution $(b_i|\, Z_{ij}) \in N({M_i}^C, {V_i}^C)$. The resulting distribution is a Gaussian function which attempts to capture the bias of each expert, as well as the precision of each expert, which together represent its skills.

We proceed next to use the observed knowledge about an expert to obtain an unbiased estimate of its score distribution at the time instant $j = n_T$. By re-applying Bayes theorem to update the distribution given in (4) one obtains that,

$$(Y_{n_T}|z_{i,1}, z_{i,2}, \ldots, z_{i,n^C}, x_{i,n_T}) \in N({M'_i}^C, {V'_i}^C) \tag{6}$$

with mean and variance

$${M'_i}^C = x_{i,n_T} + M_i^C \quad \text{and} \quad {V'_i}^C = V_i^C + \sigma_{i,n_T}^2. \tag{7}$$

Consider now the situation that m independent experts have delivered their authenticity scores on supervisor-training shots ($j = 1, 2, \ldots, {n^C}$) and the test shot n_T. Using the Bayesian updating again, the posterior distribution of b_i, given the scores at the instant $j = n_T$ and the earlier errors, is normal;

$$\begin{aligned}(Y_{n_T}|z_{1,1}, \ldots, z_{1,n^C}, x_{1,n_T}, \ldots, z_{m,1}, \ldots z_{m,n^C}, x_{m,n_T}) \\ \in N({M''_i}^C, {V''_i}^C)\end{aligned} \tag{8}$$

where

$${M''}^C = \frac{\sum_{i=1}^{m} \frac{{M'_i}^C}{{V'_i}^C}}{\sum_{i=1}^{m} \frac{1}{{V'_i}^C}} \quad \text{and} \quad {V''}^C = \left(\sum_{i=1}^{m} \frac{1}{{V'_i}^C} \right)^{-1} \tag{9}$$

However, to compute these means and variances, the score variances ${\sigma_{ij}}^2$ are needed. We suppose that these estimations are delivered by experts depending on, e.g. the quality of the current biometric sample underlying their scores. This is reasonable because not all samples have the same (good) quality, influencing the precision of the observed score x_{ij}. In case this is not practicable for various reasons, one can assume that x_{ij} has the same variance within an expert i (but allow it to vary between experts). Then, the variances of the distributions of x_{ij} need not be delivered to the supervisor, but can be estimated by the supervisor, as discussed in the following section. Before one can use the distribution $N({M''}^C, {V''}^C)$ as a supervisor, one needs to compare it with the distribution obtained by an alternative aggregation.

Assume now that we perform this training with n^I impostor samples ($Y_j = 0$) i.e. that we compute the bias distribution $N(M_i^I, V_i^I)$ when expert i evaluates impostors, and the final distribution $N({M''}^I, {V''}^I)$, with I being a label denoting "Impostor." We do not write the update formulas explicitly as these are identical

to (5,7,9) except that the training set consists of impostors. One of the two distributions $N(M''^C, V''^C)$, and $N(M''^I, V''^I)$, represents the true knowledge better than the other at the test occasion, $j = n_T$. At this point one can choose the distribution that achieves a resemblance that is most bona-fide to its role, e.g.

$$M'' = \begin{cases} M''^C, & \text{if} \quad 1 - M''^C \leq M''^I; \\ M''^I, & \text{otherwise.} \end{cases} \tag{10}$$

In other words if the client-supervisor has a mean closer to its goal (one, because $Y_j = 1$ represents client) than the impostor-supervisor's mean is to its goal (zero) then the choice falls on the distribution coming from the client-supervisor and vice-versa. An additional possibility is to reject to output a distribution in case the two competing distributions overlap more than a desired threshold. One could also think of a hypersupervisor to reconcile the two antagonist ▶ supervisor opinions.

In practice most experts deliver scores that are between 0 and 1. However, there is a formal incompatibility of this with our assumptions because the distributions of Z_{ij} would be limited to the interval $[-1, 1]$ whereas the concept we discussed earlier is based on normal distributions taking values in $]-\infty, \infty[$. This is a classical problem in statistics and is addressed typically by remapping the scores so that one works with "odds" of scores

$$X_{ij} = \log \frac{X'_{ij}}{1 - X'_{ij}} \tag{11}$$

where $X'_{ij} \in]0,1[$. It can be shown that the supervisor formula (10) and its underlying updating formulas hold for X'_{ij} as well. The only difference is in the conditional distributions which will be log normal yielding, in particular, the expected value $\exp(M'' + V''/2)$ and the variance $\exp(2M'' + 2V'') - \exp(2M'' + V'')$ for Y_{n_T}, (8).

Quality estimations for Bayesian supervisors. There are various ways to estimate the variance of a score distribution associated with a particular biometric sample on which an expert expresses an opinion of authenticity. A Bayesian supervisor expects this estimate because it works with distributions to represent the knowledge/opinion concerning the current sample as well as the past experience, not scalars. It makes most sense that this information is delivered by the expert or by considering the quality of the score. Next we discuss how these can be entered into update formulas.

One can assume that the experts give the precisions correctly except for an individual proportionality constant.

$$s_{ij} = a_i \sigma_{ij}^2 \tag{12}$$

Applying the Bayes theory again, i.e. a_i is first modeled to be a distribution rather than a scalar, then the distribution of $(a_i | (z_{i,1}, s_{i,1}), \ldots, (z_{i,n}, s_{i,n}))$ can be computed (it is a beta distribution under reasonable assumptions [4]). In turn this allows one to estimate the conditional expectation of $\frac{1}{a_i}$, yielding a Bayesian estimate of the score-error variances

$$\begin{aligned} \bar{\sigma}_{i,j^2} &= E(\sigma^2_{in_T} | s_{in_T}, (z_{i,1}, s_{i,1}), \cdots (z_{i,n}, s_{i,n})) \\ &= s_{ij} E\left(\frac{1}{a_i}\right) = s_{ij} \alpha_i = s_{ij} \frac{(G_i - D_i)}{n - 3} = \end{aligned} \tag{13}$$

with

$$\alpha_i = E\left(\frac{1}{a_i}\right) = \frac{(G_i - D_i)}{n - 3}, \qquad G_i = \sum_{j=1}^{n} \left(\frac{z_{ij}^2}{s_{ij}}\right) \quad \text{and}$$

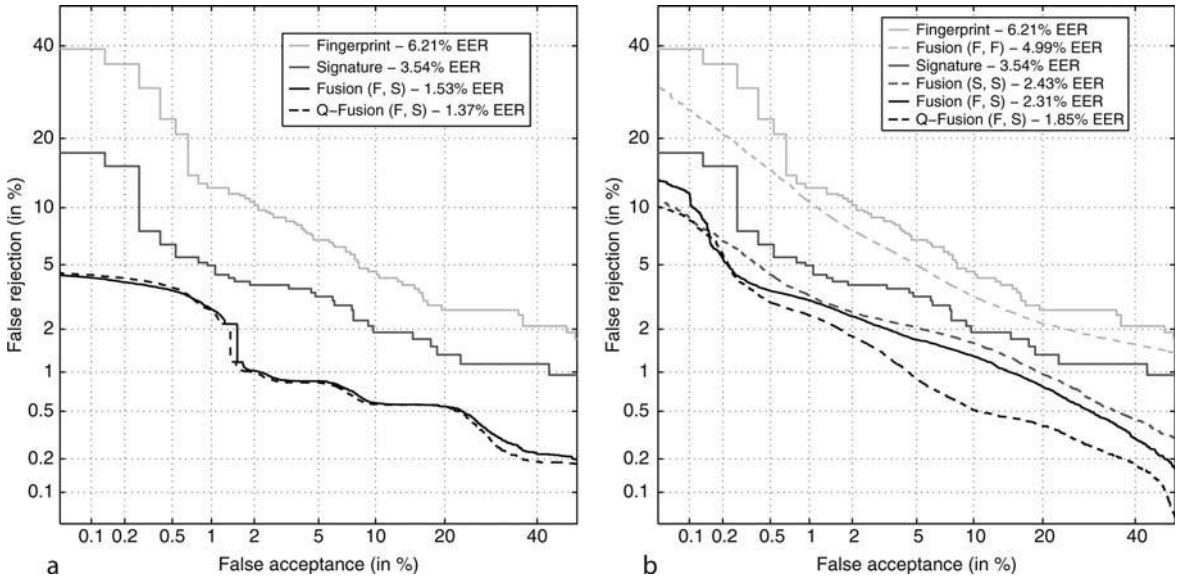
$$D_i = \left(\sum_{j=1}^{n} \left(\frac{z_{ij}}{s_{ij}}\right)\right)^2 \left(\sum_{j=1}^{n} \left(\frac{1}{s_{ij}}\right)\right)^{-1} \tag{14}$$

Note that, n will normally represent the number of biometric samples in the training set and equals to either n^C or n^I. From this result it can also be concluded that if an expert is unable to give a differentiated quality estimation then its variance estimation s_{ij} will be constant across the biometric samples it inspects and the $\bar{\sigma}_{ij}^2$ will approach gracefully to the variance of the error of the scores of the expert (not adjusted to sample quality).

The machine expert will, in practice, be allowed to deliver an empirical quality score p_{ij} because these are easier to obtain than variance estimations, s_{ij}. At this point, one can assert that these qualities are inversely proportional to the underlying standard deviations of the score distributions, yielding

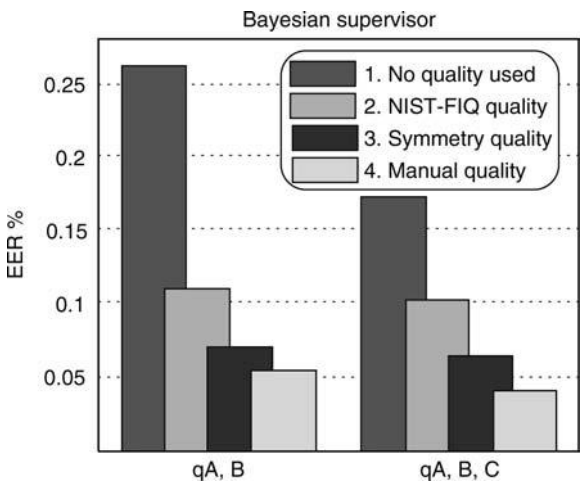
$$s_{ij} = \frac{1}{p_{ij}^2} \tag{15}$$

where p_{ij} is a quality measure of the biometric sample j as estimated by the expert i. If it is a human expert that estimates the quality p_{ij} it can be the length of the interval in which she/he is willing to place the score x_{ij}, so that even human and machine opinions can be reconciled by using the Bayesian supervisor. In Fig. 2 (a),



Multiple Experts. **Figure 2** The graphs, from [12], illustrate the recognition performance of two supervisors on the same data-set. In a probabilistic supervisor (Bayesian) and in (b) a discrimination based supervisor (SVM) were used. The used experts were common, F: fingerprint, and S: signature. In (b) there were two different fingerprint experts as well as two different signature experts. Q-Fusion represents quality based fusion.

the performance of using this Bayes supervisor in a recognition system relying on a fingerprint and a signature expert is shown. The quality scores are generated by human experts independent of the experiment. To automatically find quality scores p_{ij} for biometric samples is an emerging field of study [8–11]. The results of Bayes supervisor in connection with machine-delivered quality scores are illustrated by Fig. 3 where three fingerprint (machine) experts' opinions are weighted to yield the supervisor opinion. The experts are called A, B, and C and the quality measures used were (1) no quality used, i.e. $p_{ij} = 1$ (2) An automatic quality measure [10], (the method is publicly made available by NIST), (3) another automatic quality measure based on local symmetries [11], (4) Quality measures provided by human experts. At each experiment, one of the four quality measures is attached to the scores of A (so that this expert is called qA) in a two or three expert configurations to evaluate the effect of using sample adaptive quality measures in machine supervisors. As can be seen, using quality measures does improve the recognition performance. It is not surprising that human experts perform better in quality estimation, as this is a very complex task in which human experts still excel. However, the machine-delivered quality estimates are fairing quite well, not too far away behind human assessments of the quality. It is also worth noting that the final decisions are suggested by the machine supervisor which processed both human and machine delivered opinion parameters transparently.



Multiple Experts. **Figure 3** The graphs, adapted from [11], illustrate the recognition performance of two (qA, B) and three (qA, B, C) fingerprint experts combined by the Bayesian supervisor with automatically extracted quality measures (attached to qA).

Discrimination Functions Based Fusion

Discrimination functions are frequently used in pattern classification and can also be used to fuse decisions of biometric experts. Discriminative statistics differs from Bayesian approach in that distributions of random variables are not necessary for decision making. Instead, modeling the decision boundaries is the focus of attention. Here we exemplify the use of this approach in fusion by Support Vector Machines, SVMs [3].

Assume that we are given a set of observations

$$\{\mathbf{x}_1, y_1\}, \{\mathbf{x}_2, y_2\} \cdots \{\mathbf{x}_n, y_n\} \tag{16}$$

where $\mathbf{x}_j = (x_{1,j}, \ldots, x_{1,j})^T$ is a feature vector of dimension m, and y_j is the class-label of the latter (relative to the classes, "client" and "impostor"), respectively. Assume further that the two classes are separable by a hyperplane. Then there is an optimal hyperplane in a high dimensional space to which $\mathbf{x}$ is mapped. For simplicity we assume that the mapping is the (trivial) identity transformation but other tranformations using e.g. polynomials, or radial basis functions, can be used with little impact on the discussion that follows next [3], provided that appropriate kernel functions are used whenever scalar products are utilized in computations. The separation hyperplane

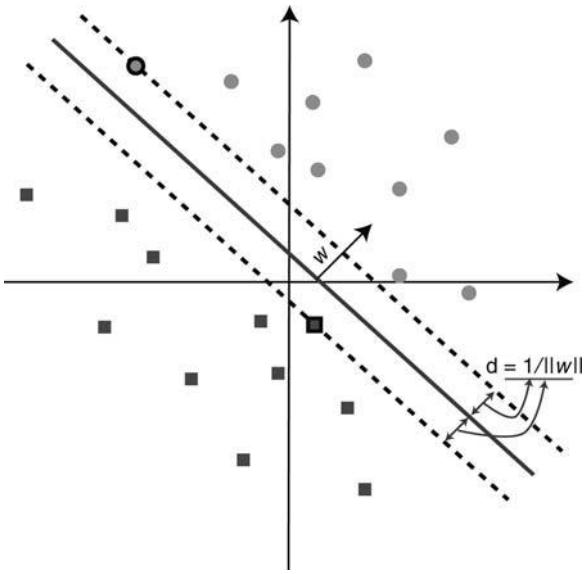
$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b = 0 \tag{17}$$

can be made to have maximal distance d_{max} to samples belonging to the two classes, Fig. 4. The equation can be multiplied by a nonzero constant such that $\|\mathbf{w}\| = 1/d_{max}$. We can (using this freedom) represent thereby the two class-labels as $+1$ and -1, to follow the convention of SVM litterature. Then, we have

$$\begin{cases} f(\mathbf{x}_j) \geq 1 & \text{if } y_j = 1 \\ f(\mathbf{x}_j) \leq -1 & \text{if } y_j = -1 \end{cases} \tag{18}$$

Equivalently, the distance is maximized if $\frac{1}{2}\|\mathbf{w}\|^2$ is minimized under the constraints given by (18). If we know $\mathbf{w}$ and b, the function f will be a discrimination function, e.g. $f(\mathbf{x}) \geq 0$ prompts for a decision $y = 1$. The parameters $\mathbf{w}$ and b can be found by solving a quadratic problem with linear constraints.

In case the classes are not separable by a hyperplane, slack variables ξ_j are introduced so as to allow a classification that makes an error, but that this error is the smallest on the training/observation set. The



Multiple Experts. Figure 4 Illustration of two classes (circle:client and square:impostor) that are separable by a hyperplane with direction $\mathbf{w}$. The support vectors that define the separation hyperplane are represented by the outlined square and the circle on dashed hyperplanes. The width of the separation zone is $2/\|\mathbf{w}\|$ which is maximized by SVM.

corresponding problem is still a quadratic optimization problem

$$\min \frac{1}{2}\|\mathbf{w}\|^2 + \sum_j C\xi_j \tag{19}$$

subject to the constraints

$$\begin{cases} f(\mathbf{x}_j) \geq 1 & \text{if } y_j = 1 - \xi_j \\ f(\mathbf{x}_j) \leq -1 & \text{if } y_j = -1 + \xi_j \end{cases} \tag{20}$$

The constant C assures that there is a limit on the amount of change the training vectors can introduce to the solution.

The SVM formulation allows one to construct a supervisor that is able to assign a class label y_j to the score vector of m experts $\mathbf{x}_j = (x_{1,j}, x_{2,j} \ldots x_{m,j})^T$ (at $j = n_T$). However, such a supervisor would not be quality adaptive yet, because the contribution of each sample to the total cost function would be uniform due to C's being a constant. This can be changed such that the cost depends on the quality of the biometric sample by means of a heuristically chosen function. In the following section we follow the description of [8] to obtain such a sample adaptation and a final supervisor.

Let the original quality measure delivered by the expert i be $p_{ij} \in [0, p_{max}]$ and

$$q_{ij} = \sqrt{p_{ij} \cdot \bar{p}_i} \tag{21}$$

where $\bar{p}_i$ is the average quality measure of the signals that expert i delivered for the training samples, which is also in the range of $[0, p_{max}]$. Then one can train m SVMs, each having its own discrimination function f_i, by defining the cost coefficients for the respective function as follows.

$$C_{ij} = C\left(\frac{\prod_{i' \neq i} q_{i',j}}{p_{max}^{m-1}}\right)^{\gamma_1} \tag{22}$$

The coefficient C_{ij} represents now the cost of not influencing the biometric sample j by expert i and it is measured as the product of the quality measures of other experts (excluding expert i) on the current sample j. The training samples of f_i are $\mathbf{x}_j^{\mathrm{i}}$, $j: 1 \ldots n$, which equals to $\mathbf{x}_j$ except that its component corresponding to expert i has been removed.

$$\mathbf{x}_j^{\mathrm{i}} = (x_{1,j}, \cdots, x_{i-1,j}, x_{i+1,j}, \cdots x_{m,j})^T \tag{23}$$

Here the use i superscript is in the sense of label, not exponent, signifies that the data of expert i is lacking. The exponent γ_1 is an empirically chosen constant the purpose of which is to adjust the overall influence of quality based discrimination on the final decision. In a similar fashion an additional discrimination function f_0 can be computed, except that the cost coefficients are now defined as

$$C_j = C\left(\frac{\prod_{i=1}^{m} q_{i,j}}{p_{max}^{m}}\right)^{\gamma_2} \tag{24}$$

This represents the alternative cost of using all individual quality measures including those delivered by expert i. The discrimination function f_0 is obtained by an SVM training on full length expert score vectors $\mathbf{x}_j$, $j: 1 \ldots n$, as opposed to f_i, $i: 1, \ldots, m$ which trains on $\mathbf{x}_j^{\mathrm{i}}$, lacking the opinion of expert i. When the system is operational at time $j = n_T$, the m quality scores q_{i,n_T} as well as m expert scores x_{i,n_T} are available. The quality measures q_{i,n_T}, as well as the corresponding scores q_{i,n_T} and the discrimination functions f_i, $i: 1, \ldots, m$, are re-indexed such that $q_{1,n_T} \leq \cdots \leq q_{m,n_T}$. A final supervisor can then be obtained by aggregating f_0 with $f_1, \ldots f_m$ as follows:

$$f_Q = \beta_1 \sum_{i=1}^{m-1} \frac{\beta_i}{\sum_{i'=1}^{m-1} \beta_{i'}} f_i(\mathbf{x}_{n_T}^i) + (1 - \beta_1) f_0(x_{n_T}) \tag{25}$$

where

$$\beta_i = \left(\frac{q_{m,n_T} - q_{i,n_T}}{p_{\max}}\right)^{\alpha_2} \tag{26}$$

The results of this supervisor is shown in Fig. 2 (2) where fingerprint and signature traits are fused using human expert opinions. Again, one can conclude that skill and sample adaptation do help to improve the recognition performance.

Related Entries

▶ Quality Measures

References

1. Ross, A., Jain, A.K.: Information fusion in biometrics. Pattern Recogn. Lett. **24**(13), 2115–2125, (2003)
2. Lindley, D.V.: Making decisions. Wiley, London (1990)
3. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**, 273–297 (1995)
4. Bigun, E., Bigun, J., Duc, B., Fischer, S.: Expert conciliation for multi modal person authentication systems by bayesian statistics. In: Bigun, J. Chollet, G. Borgefors G. Audio and Video based Person Authentication – AVBPA97, pp. 291–300. Springer, Heidelberg (1997)
5. Kittler, J., Li, Y., Matas, J., Sanchez, M.R.: Combining evidence in multimodal personal identity recognition systems. In: Bigun, J. Chollet, G. Borgefors G. Audio and Video based Person Authentication - AVBPA97, pp. 327–334. Springer, Heidelberg (1997)
6. Ben-Yacoub, S., Abdeljaoued, Y., Mayoraz, E. Fusion of face and speech data for person identity verification. IEEE Trans. Neural Networks **10**(5), 1065 (1999)
7. Gutschoven, B., Verlinde, P.: Multi-modal identity verification using support vector machines (SVM). In: Proceedings of the Third International Conference on Information Fusion. IEEE Press (2000). http://citeseer.ist.psu.edu/303146.html;ftp://ftp.elec.rma.ac.be/user/verlinde/pub/fusion_2000.ps.gz
8. Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J., Bigun, J.: Discriminative multimodal biometric authentication based on quality measures. Pattern Recogn. **38**, 777–779 (2005)
9. Lim, E., Toh, K.A., Saganthan, P.N., Jiang, X.D., Yan, W.Y.: Fingerprint image quality analysis. In: International Conference on Image Processing, pp. II: 1241–1244 (2004). URL http: //dx.doi.org/10.1109/ICIP.2004.1419530
10. Tabassi, E., Wilson, C.L.: A novel approach to fingerprint image quality. In: ICIP (2), pp. 37–40 (2005). URL http://dx.doi.org/10.1109/ICIP.2005.1529985

11. Fronthaler, H., Kollreider, K., Bigun, J., Fierrez, J., Alonso-Fernandez, F., Ortega-Garcia, J.: Automatic image quality estimation and its application to multi-algorithm verification. IEEE Transactions on Information Forensics and Security p. **3**(2): 331–338 (2008)
12. Fierrez-Aguillar, J.: Adapted fusion schemes for multimodal biometric authentication. Phd thesis, Universidad Politecnica de Madrid, ETSIT (2006)

Multiple View Geometry

Multiple view geometry is the scientific field which studies the geometric relations between images of the same objects taken from different views. It relies heavily on affine- and projective-geometry.

► Palm Vein Image Sensor

Multispectral and Hyperspectral Biometrics

Besma Roui-Abidi, Mongi Abidi
The University of Tennessee, Knoxville, TN, USA

Synonyms

Imaging spectroscopy; Multiband biometrics; Fusion, Wavelet-Based

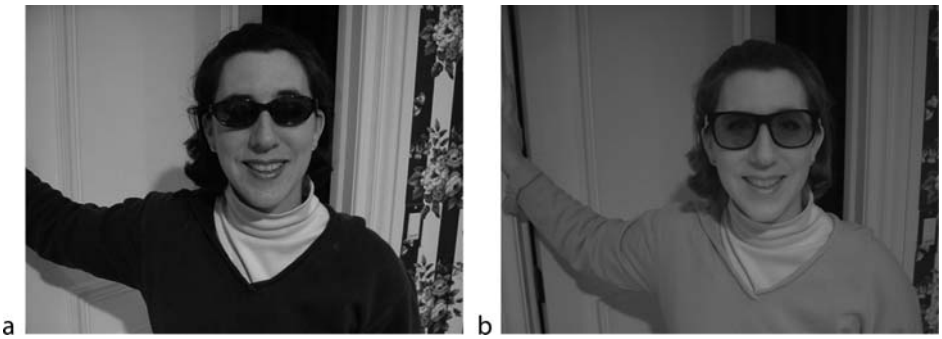
Definition

Multispectral biometrics are based on data consisting of four to ten separate images of the same biometric trait and representing sensors' responses in different wavelengths of the electromagnetic spectrum. In contrast to conventional images, which generally represent an integrated response of a single sensor over a wide range of bands in the same spectral zone, multispectral data usually refers to multiple separate sensor(s) responses in relatively narrow spectral bands. The word *multispectral* was first used in space-based imaging to denote data acquired in the visible and infrared spectra. In biometrics, the word *multispectral* has been used to describe responses in multiple narrow bands either all in the visible, or all in the infrared, or a mixture of both. Even though the words *hyperspectral* and *multispectral* have often been used interchangeably, hyperspectral imaging usually refers to cases where the number of bands are higher than 10 and when these bands encompass more than one region of the electromagnetic spectrum, such as the visible and the infrared. Without lack of generality, the word *multispectral* will be used in the reminder of this article to refer to any biometric work involving more than the three customary red, green, and blue channels.

Introduction

Multispectral imaging provides data in the form of a three dimensional image cube, with two spatial dimensions (horizontal and vertical coordinates) and one spectral dimension. Also called imaging spectroscopy, this type of data details the spectral content of each pixel in the 2D image, therefore providing more than what the human eye can capture with its receptors in the red, green, and blue. The separation of a pixel's content into information within multiple very narrow wavelengths allows the material imaged to be identified based on its spectral characteristics in addition to its spatial characteristics. The high number and different characteristics of the bands make it easier to differentiate between objects that would look similar in regular intensity images or even in conventional color images. In other words, a multispectral dataset is a higher resolution image in the spectral dimension that makes it possible to resolve information and details non resolvable in conventional images.

The higher dimensionality of multispectral data presents a desirable feature for biometric systems, i.e., the uniqueness of a material based on its spectral characteristics, including material of a person's skin, iris, and vasculature. A conventional sensor's response to the combination of material and illumination is not always most informative in the visible domain. For instance, conditions such as night time and low intensity light are better dealt within the infrared domain, while the visible spectrum generally reveals better information in uniform well lit situations. Concealment and disguise is another issue that is usually hard to deal with in the visible domain alone. Figure 1 shows two images of a person wearing sunglasses, one in the visible light and the other in the near infrared (NIR).



Multispectral and Hyperspectral Biometrics. Figure 1 Visible and NIR images of a subject wearing sun glasses. Images are acquired using an IR-Enabled Sony DSC-S30 Digital Camera & X-Nite780 nm filter under incandescent lighting http://www.maxmax.com/aXRayIRExamplesSecurity.htm.

The eyes, which are fundamental features for any face recognition algorithm, appear completely blocked out in the visible image, while visible in the NIR image.

Versatility, usability, and security are some of the required characteristics of any biometric system. Such system must have the capability to acquire and process biometric data at different times of day and night, in a variety of weather and environmental conditions, and be resistant to spoofing attempts. Multispectral biometrics is one of the few technologies shown to solve many of the aforementioned issues. Multispectral analysis has been used to improve recognition rates and detect spoofing attempts for various biometric modalities, including face [1], iris [2], fingerprint [3], and vasculature [4]. Wavelengths covering the visible spectrum all the way to the long wave infrared have been used in the analysis of Biometric data. Figure 2 displays the distribution and values of wavelengths along the electromagnetic spectrum, with emphasis on the most used bands in biometrics. The visible spectrum comprises wavelengths between 400 and 750 nm. Bands of lengths from 750 to 1,400 nm represent the NIR, while wavelengths between 1.4 and 3 μm are called short wave infrared (SWIR). Bands from 3 to 8 μm are said to be in the mid wave infrared (MWIR) and bands between 8 and 15 μm are called long wave infrared (LWIR). Mid wave and long wave are also referred to as thermal infrared.

Multispectral Biometric Data Acquisition

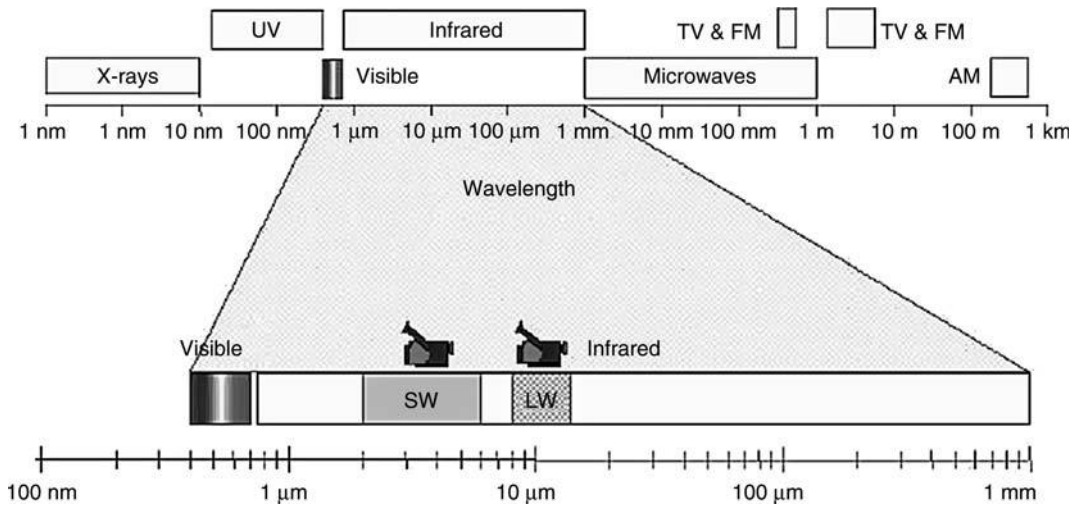
Most multispectral instruments use a 2D detector array and are scanned over time to acquire the third dimension of the cube. Some common techniques for acquiring multispectral data are: (1) Spatially scanning split spectrometer, (2) wavelet tunable spectral filter, (3) two-dimensional Fourier transform imager, or (4) a combination of sensors with responses in different regions of the spectrum. A number of instruments were used in Biometrics to acquire multispectral data of face, iris, fingerprint, and vasculature. Separate but aligned/calibrated visible and infrared cameras were used to obtain registered face images [5, 6]. Liquid crystal tunable filters attached to monochromatic video cameras were also deployed to acquire multiple spectral bands (25 bands between 400 and 800 nm [7] and 31 bands between 700 and 1,000 nm [8]) of face images over time. Tunable filters allow the band to be changed electronically and continuously with selectable resolution and interval. A spectropolarimetric camera was used for hyperspectral face data acquisition in the range of 450–1,100 nm [9]. Iris images in the visible and NIR were obtained using a multispectral camera consisting of three charge coupled devices and three band-pass prisms, resulting into a 4 band output [2]. Multispectral fingerprint data has been collected using a single placement of the finger on a sensor that combines a conventional optical fingerprint reader and multiple illuminators with different wavelengths, light orientations, and polarization conditions to yield information on both the surface and subsurface region of the finger's skin [3].

Multispectral Face

Changes in conditions between gallery and probe images have often resulted in major performance degradations of face recognition algorithms. Such conditions

include illumination, pose, time lapse, and expression [10]. A number of studies were conducted in recent years using multispectral face data in an attempt to establish invariance of face biometric systems to changes in these conditions. Researchers and scientists addressed various multispectral aspects of face recognition. These included restricting the bands to certain zones of the spectrum, such as using just visible bands or only infrared wavelengths, as well as combining spectral bands from the visible, NIR, and/or thermal infrared. Figure 3 displays a sample set of face images acquired in different bands of the electromagnetic spectrum.

A study conducted at the University of Tennessee used 25 bands in the visible spectrum to analyze the effect of change in illumination conditions, especially outdoor versus indoor lighting, on face recognition performance. Chang et al. analyzed the multispectral face data in light of the measured spectral distributions of the illuminants and addressed spectral band selection as well as the fusion of multispectral data to improve invariance to illumination [7, 11]. Fusion methods such as averaging, Principal Components Analysis, ▶ wavelet analysis, physics based, and illumination adjustment were implemented and tested. Rank-based decision fusion was also tested on various bands



Multispectral and Hyperspectral Biometrics. **Figure 2** Bands of the electromagnetic spectrum and their wavelengths www.itcnewsletter.com [Courtesy: Infrared Training Center].

Multispectral and Hyperspectral Biometrics. **Figure 3** A set of multispectral face images acquired in 8 single narrow bands of the visible spectrum, followed by the three R, G, B components from a conventional color camera, the composite color image, a NIR, and a thermal infrared image of the same subject- (Courtesy of Imaging, Robotics, and Intelligent Systems lab, University of Tennessee).

and shown to considerably improve recognition performance. The authors of the study demonstrated that fused face data outperforms conventional images especially under severe changes in illumination conditions and long time lapses between gallery and probe images.

Using a database of 200 face images spanning 31 bands of the NIR spectrum, Pan et al. demonstrated that the subsurface information provided in these bands can be unique to each individual, is relatively stable over time, and invariant to pose and expression variations [8, 12].

Other studies on face biometrics involved a combination of spectral data from different spectral zones. One of the first studies combining visible and infrared face images was conducted by Sokolinsky et al. [5], who were able to show an increase in performance by fusing visible and infrared spectral bands. Kong et al. demonstrated that wavelet-based fusion of visible and thermal infrared imagery considerably improved face recognition under a variety of illumination directions and intensities [1].

Buddharaju et al. combined data in the visible with data in the thermal spectra and used the physiological characteristics of the thermal face vasculature to improve recognition rates [4] of face images with varying expressions. A score level fusion approach was used. The authors also studied the effect of change in temperature and sweating on the performance of the algorithm and found that extreme conditions resulted in non linear variations in thermal data and therefore deterioration in performance. Kakadiaris et al. combined 3D face reconstructed from multiple visible sensors with the texture and vasculature emanating from a calibrated infrared camera to cancel out the effect of expression changes on recognition performance [6]. Skin temperature was encoded in the metadata records used for recognition and the algorithm showed a minimal drop in performance in the presence of changes in facial expression compared to the baseline algorithms used on the Face Recognition Grand Challenge database.

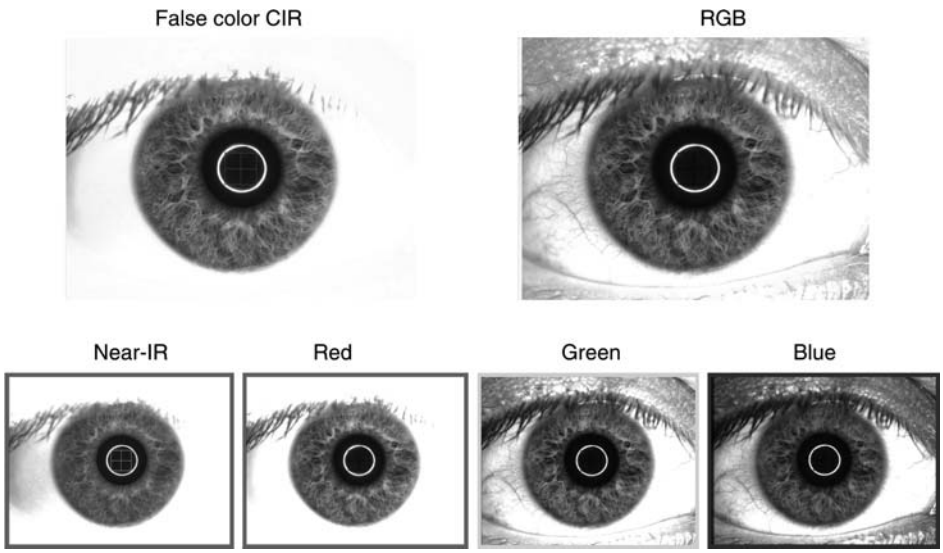
Multispectral Iris

Multispectral iris processing and recognition is a relatively new area of biometrics. Early work by Imai was conducted for a purely clinical application. It describes experiments for the spectral characterization of the human iris using an ophthalmic microscope, a digital camera, and a spectroradiometer under halogen light. The spectral bands studied were all in the visible range from 400 to 700 nm and the study showed that the recovered iris' spectra of different subjects presented clear distinctive differences. No database was built and no biometric processing was conducted within the study [13]. Boyer et al. acquired multispectral data of the iris in the visible red, green, and blue and in the NIR in an effort to demonstrate that different characteristics of the iris will show in different bands based on the color of the eye [2]. Figure 4 shows an example of a blue eye iris in the various bands. The authors used an adaptive color histogram equalization technique to enhance the iris structure and reveal the needed information in the individual channels. They then evaluated the performance of iris recognition in the various bands using a database of 24 subjects and 5 samples per subject. A before and after performance evaluation showed a substantial improvement in the blue channel. The authors then conducted a series of experiments of cross matching between individual channels and showed that the performance degrades with the increase in distance between the matched channels. This indicates a decrease in correlation between the information revealed in the more distant bands. The study went on to show potential improvements to iris segmentation using color based clustering.

Park and Kang processed infrared multispectral iris data using a gradient-based fusion scheme. The purpose of the study was to detect spoofing attempts by using the specific complementary information contained in the various bands of a real iris. Images with no spectral variation (fake irises) resulted in an erroneous fused image that could not be matched against actual irises [14].

Multispectral Fingerprint

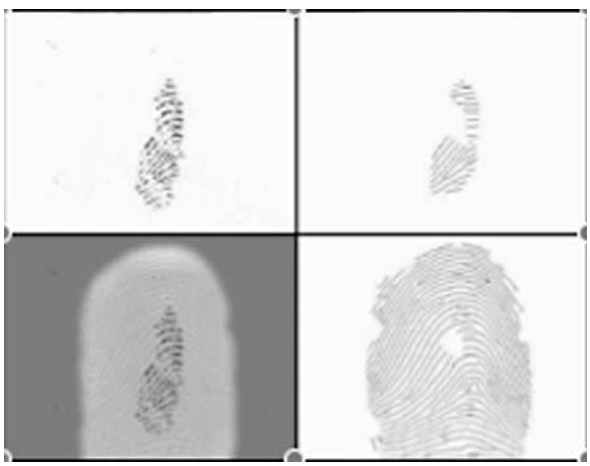
Fingerprint sensors are widely used in law enforcement, cyber, and physical security applications. However, the performance of these sensors is easily degraded by physiological and environmental conditions, such as dry or wet skin, heat or cold, and ambient lighting conditions. These factors can result in failure to enroll, a high rate of false rejection, and dissatisfied users. Multispectral fingerprint Biometrics is a relatively new technology aiming to solve some of



Multispectral and Hyperspectral Biometrics. **Figure 4** An example of a blue iris showing the false color NIR, the RGB composite, and each individual channel [2].

the aforementioned problems. Multispectral fingerprint was first introduced by Rowe and Nixon [3], who designed and tested a sensor able to acquire fingerprint data in five visible spectral bands (445, 500, 574, 610, and 660 nm). The purpose of the effort was to improve the usability of fingerprint biometric systems by improving the robustness of the sampling process to physiological and environmental conditions such as dry skin, rain, heat, lighting conditions, cold, etc. The authors analyzed fingerprint recognition on a number of samples from a baseline database of 4,105 samples, a cold finger study of 300 samples, and a wet finger study of 1,186 samples using multiple sensors of the same design. A dramatic improvement in Equal Error Rate was achieved by using the new multispectral acquisition technique compared to the customary total internal reflectance technique (IRT). An example of multispectral finger data and an IRT optical sample of a fingerprint are shown in Fig. 5. The subject was asked to apply too little pressure on the sensors. In addition to a better and more successful enrollment, the new multispectral sensor also provides information on the subsurface of the finger and other attributes that makes spoofing by using an artificial rubber or other fake fingers difficult to achieve [15].

A different form of multispectral analysis called Fourier Transform Infrared (FTIR) chemical imaging was used in forensic studies to recover and identify latent fingerprints from traditionally hard to analyze



Multispectral and Hyperspectral Biometrics. **Figure 5** Fingerprint scans (*left*) and their enhanced versions (*right*) using a conventional IRT optical sensor (*top*) and a multispectral sensor (*bottom*). The subject was asked to apply very little pressure on the sensor [15].

backgrounds [16]. In this approach, the infrared spectrum of the sample is collected by passing a beam of infrared light through the sample. A systematic methodology is used for each surface by optimizing the spectral resolution, number of scans, and pixel aggregation. Examination of the transmitted light reveals how much energy was absorbed at each wavelength. A Fourier transform instrument measures all

wavelengths at once. From this, a transmittance or absorbance spectrum is produced. Analysis of these absorption characteristics reveals details about the molecular structure of the sample. The use of this approach was shown to improve sample analysis from polymer bank notes and aluminum drink cans. Crane et al. used various processing techniques on FTIR images of fingerprints on a number of challenging porous and nonporous substrates to extract the ridge patterns. Techniques used include basic infrared spectroscopic band intensities, addition and subtraction of band intensity measurements, principal components analysis, and calculation of second derivatives band intensities. Trace evidence within the fingerprints were also recovered and identified.

Related Entries

- ▶ Biometric System
- ▶ Face Recognition
- ▶ Fingerprint Recognition
- ▶ Iris Recognition
- ▶ Liveness and Anti Spoofing
- ▶ Near Infrared Based Face Recognition
- ▶ Skin Spectroscopy

References

1. Kong, S.G., Heo, J., Boughorbel, F., Zheng, Y., Abidi, B.R., Koschan, A., Abidi, M.A.: Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition. Int. J. Comput. Vis. **71**(2), 223–253 (2007)
2. Boyce, C., Ross, A., Monaco, M., Hornak, L., Li, X.: Multispectral iris analysis: a preliminary study. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition Workshop on Biometrics (CVPRW), New York. IEEE publisher, New York (2006)
3. Rowe, R., Nixon, K.A., Corcoran, S.P.: Multi spectral fingerprint biometrics. In: Proceedings of IEEE Workshop on Information Assurance and Security, United States Military Academy, West Point, NY, pp. 14–20 (2005)
4. Buddharaju, P., Pavlidis, I.: Multi-spectral face recognition – fusion of visual imagery with physiological information. In: Hammoud, R., Abidi, B., Abidi, M. (eds.) Face Biometrics for Personal Identification. Springer, Berlin, pp. 91–108 (2007)
5. Socolinsky, D.A., Wolff, L.B., Neuheisel, J.D., Eveland, C.K.: Illumination invariant face recognition using thermal infrared imagery. In: Proceedings of the IEEE ICPR, CVPR, Kauai, HI, pp. 527–534 (2001)
6. Kakadiaris, I.A., Passalis, G., Toderici, G., Lu, Y., Karampatziakis, N., Murtuza, N., Theoharis, T.: Expression-invariant multispectral face recognition: you can smile now! In: Flynn, P.J., Sharath Pankanti. (eds.) Proceedings of SPIE. Biometric Technology for Human Identification III, vol. 6202, pp. 620 204.1–620 240.7 (2006)
7. Chang, H., Harishwaran, H., Yi, M., Koschan, A., Abidi, B., Abidi, M.: An indoor and outdoor, multimodal, multispectral and multi-illuminant database for face recognition. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Biometrics Work Shop, New York, NY (2006)
8. Pan, Z., Healey, Z., Prasad, M., Tromberg, B.: Face recognition in hyperspectral images. IEEE Trans. Pattern Anal. Mach. Intell. **25**(12), 1552–1560 (2003)
9. Denes, L.J., Metes, P., Liu, Y.: Hyperspectral face database. Technical report CMU-RI-TR-02-25, Robotics Institute, Carnegie Mellon University (2002)
10. Phillips, P.J., Grother, P., Micheals, R.J., Blackburn, D.M., Tabassi, E., Bone, J.M.: FRVT 2002: Evaluation Report (2003). http://www.frvt.org/FRVT2002/documents.htm
11. Chang, H., Koschan, A., Abidi, B., Abidi, M.: Physics-based fusion of multispectral data for improved face recognition. In: Proceedings of the International Conference on Pattern Recognition, Hong Kong (2006)
12. Pan Z., Healey G., Prasad M., Tromberg B.: Multiband and spectral eigenfaces for face recognition in hyperspectral images. In: Proceedings of SPIE, San Jose, CA, vol. 5779, pp. 144–151 (2005)
13. Imai, F.H.: Preliminary experiment for spectral reflectance estimation of human iris using a digital camera, Munsell Color Science Laboratory Technical Report (2002)
14. Park, J.H., Kang, M.G.: Iris recognition against counterfeit attack using gradient based fusion of multi-spectral images. Advances in Biometric Person Authentication, LNCS vol. 3781. Springer, Berlin, pp. 1611–3349 (2005)
15. Nixon, K.A., Rowe, R.K.: Multispectral fingerprint imaging for spoof detection. In: Jain, A.K., Ratha, N.K. (ed.) Proceedings of SPIE. Biometric Technology for Human Identification II, vol. 5779, pp. 214–225 (2005)
16. Crane, N.J., Bartick, E.G., Perlman, R.S., Huffman, S.: Infrared Spectroscopic Imaging for Noninvasive Detection of Latent Fingerprints. Forensic Sci. **52**(1), 48–53 (2007)

Multistage Matching

Multistage matching is a technique used in order to simultaneously achieve high accuracy and high speed during the matching stage: A fast initial algorithm is used to compare the query to each template of the

database, and a decision is made to disregard or keep the template for the next stage. A slower, more accurate, algorithm is then applied to the surviving templates, eliminating the additional database entries. This process is repeated until the last stage, where the final decision will be made. In multistage matching, the slowest algorithms are applied only to a few numbers of templates and, hence, have a small impact on the overall matching speed.

► Biometric Algorithms

Mutual Authentication

Mutual authentication or two-way authentication is a process in which two entities in communication authenticates each other before any application data is transferred. This is typically achieved by exchange of digital certificates issued by trusted entities. Mutual authentication helps in eliminating the man-in-the-middle attack, where an adversary establishes independent links with both the victims and relays messages between them. The victims are led to believe that they are in direct communication, while infact, the entire communication between them is controlled by the adversary.

► Security Issues, System Design

Mutual information

The mutual information $I(x1;\ x2)$ of two random variables $x1$ and $x2$ is the relative entropy between the joint distribution $p(x1,\ x2)$ and the product distribution $p(x1)\mathrm{p}(x2)$,

$$I(x1;x2) = \frac{p(x1,x2)\log(p(x1,x2))}{p(x1)\,p(x2)}\mathrm{d}x1\mathrm{d}x$$

► Independent Component Analysis

N

NAP-SVM

The main goal of the SVM Nuisance Attribute Projection (NAP) method is to reduce the impact of channel variations (called also session variability). It uses an appropriate projection matrix P in the SVM supervector space (a supervector is obtained by concatenating the Gaussian means) to remove the subspace that contains the session variability.

$$s' = P \times s, \tag{1}$$

where, s is a GMM supervector. The projection matrix can be written as follow:

$$P = (I - VV^t), \tag{2}$$

where, $V = [v_1, ..., v_k]$ is a rectangular matrix of low rank whose columns are orthonormal. The vectors v_k are obtained from the k eigenvectors having the k largest eigenvalues of the following covariance matrix:

$$\frac{1}{S}\sum_{s=1}^{S}\frac{1}{n_s}\sum_{i=1}^{n_s}(\bar{\mu}_i^s - \bar{\mu}_i^s)(\bar{\mu}_i^s - \bar{\mu}_i^s)^t \tag{3}$$

where $\bar{\mu}_i^s$ represents the GMM supervector of the ith session of the sth speaker. S is the number of speaker in V -Matrix training data. n_s is the number of different sessions belonging to the sth speaker. $\bar{\mu}_i^s$ is the mean GMM supervector obtained overall the sessions belonging to the sth speaker:

$$\bar{\mu}_i^s = \frac{1}{n_s}\sum_{i=1}^{n_s}\bar{\mu}_i^s. \tag{4}$$

► Session Effects on Speaker Modeling

National Institute for Standards and Technology

► Fingerprint, Forensic Evidence of

Natural Gradient

When a parameter space has a certain underlying structure, the ordinary gradient (partial derivative) of a function does not represent its steepest direction. Riemannian space is a curved manifold where, there is no orthonormal linear coordinates. The steepest descent direction in a Riemannian space is given by the ordinary gradient pre-multiplied by the inverse of Riemannian metric. Such direction is referred to as *natural gradient*.

► Independent Component Analysis

Near Field Communication

Synonym

NFC

Definition

A method of wireless communication that uses magnetic field induction to send data over very short distances (less than 20 cm). NFC is intended primarily for deployment in mobile handsets.

► Transportable Asset Protection

© 2009 Springer Science+Business Media, LLC

Near Infrared (NIR)

Electromagnetic radiation, identical to visible light, except at wavelengths longer than red light; the near infrared band (IR-A) extends from ∼700 to 1,400 nm, while the whole infrared region ranges from ∼700 to 3,000 nm.

- ▶ Face Recognition, Near-infrared
- ▶ Finger Vein
- ▶ Iris Databases
- ▶ Iris Device
- ▶ Palm Vein

Near-infrared Image Based Face Recognition

▶ Face Recognition, Near-Infrared

NEXUS

▶ Iris Recognition at Airports and Border-Crossings

NIST SREs (Speaker Recognition Evaluations)

Evaluations of speaker recognition systems coordinated by the National Institute of Standards and Technology (NIST) in Gaithersburg, MD, USA, 1996–2008.

▶ Speaker Databases and Evaluation

Noisy Iris Challenge Evaluation – Part I (NICE.I)

The Noisy Iris Challenge Evaluation – Part I (NICE.I) began in 2007 by the University of Beira Interior. The NICE.I contest focuses on the development of new iris segmentation and noise detection techniques unlike similar contest which focus more on iris recognition performance. The iris database used for the contest, UBIRIS.v2, consists of very noisy iris images to simulate less constraining image capturing conditions.

▶ Iris Databases

Nominal Identity

Nominal identity represents the name and in certain cases the other abstract concepts associated with a given individual (e.g. the Senator, the Principal, the actor). Such an identity is malleable and distinguished from the less mutable biometric identity, which is typically predicated on physiological or behavioral characteristics that do not change much over time.

▶ Fraud Reduction, Applications

Non-ideal Iris

Non-ideal iris is defined as dealing with off-angle, occluded, blurred, noisy images of iris.

▶ Iris Image Quality

Non-linear Dimension Reduction Methods

▶ Non-linear Techniques for Dimension Reduction

Non-linear Techniques for Dimension Reduction

Jian Yang, Zhong Jin, Jingyu Yang
School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, Peoples Republic of China

Synonyms

Non-linear dimension reduction methods

Definition

Dimension reduction refers to the problem of constructing a meaningful low-dimensional representation of high-dimensional data. A dimension reduction technique is generally associated with a map from a high-dimensional input space to a low-dimensional output space. If the associated map is non-linear, the dimension reduction technique is known as a non-linear dimension reduction technique.

Introduction

Dimension reduction is the construction of a meaningful low-dimensional representation of high-dimensional data. Since there are large volumes of high-dimensional data (such as climate patterns, stellar spectra, or gene distributions) in numerous real-world applications, dimension reduction is a fundamental problem in many scientific fields. From the perspective of pattern recognition, dimension reduction is an effective means of avoiding the "curse of dimensionality" and improving the computational efficiency of pattern matching.

Researchers have developed many useful dimension reduction techniques. These techniques can be broadly categorized into two classes: linear and non-linear. Linear dimension reduction seeks to find a meaningful low-dimensional subspace in a high-dimensional input space. This subspace can provide a compact representation of high-dimensional data when the structure of data embedded in the input space is linear. The principal component analysis (PCA) and Fisher linear discriminant analysis (LDA or FLD) are two well-known linear subspace learning methods which have been extensively used in pattern recognition and computer vision areas and are the most popular techniques for face recognition and other biometrics.

Linear models, however, may fail to discover essential data structures that are non-linear. A number of non-linear dimension reduction techniques have been developed to address this problem, with two in particular attracting wide attention: ▶ kernel-based techniques and ▶ manifold learning related techniques. The basic idea of kernel-based techniques is to implicitly map observed patterns into potentially much higher dimensional feature vectors by using non-linear mapping determined by a kernel. This makes it possible for the non-linear structure of data in observation space to become linear in feature space, allowing the use of linear techniques to deal with the data. The representative techniques are kernel principal component analysis (KPCA) [1] and kernel Fisher discriminant analysis (KFD) [2, 3]. Both have proven to be effective in many real-world applications.

In contrast with kernel-based techniques, the motivation of manifold learning is straightforward, as it seeks to directly find the intrinsic low-dimensional non-linear data structures hidden in observation space. Over the past few years many manifold learning algorithms for discovering intrinsic low-dimensional embedding of data have been proposed. Among the most well-known are isometric feature mapping (ISOMAP) [4], local linear embedding (LLE) [5], and Laplacian Eigenmap [6]. Some experiments demonstrated that these methods can find perceptually meaningful embeddings for face or digit images. They also yielded impressive results on other artificial and real-world data sets. Recently, Yan et al. [7] proposed a general dimension reduction framework called *graph embedding*. LLE, ISOMAP, and Laplacian Eigenmap can all be reformulated as a unified model in this framework.

Kernel-Based Non-Linear Dimension Reduction Techniques

Over the last 10 years, kernel-based dimension reduction techniques, represented by kernel principal component analysis (KPCA), and kernel Fisher discriminant analysis (KFD), have been extensively applied to biometrics and have been proved to be effective. The basic idea of KPCA and KFD is as follows.

By virtue of a non-linear mapping Φ, the *input data space* $\mathbb{R}^n$ is mapped into the *feature space* H:

$$\frac{\Phi : \mathbb{R}^n \to \mathrm{H}}{x \mapsto \Phi(x)} \tag{1}$$

As a result, a pattern in the original *input space* $\mathbb{R}^n$ is mapped into a potentially much higher dimensional feature vector in the *feature space* H. KPCA is to perform PCA in the feature space, while KFD is to perform LDA in such a space.

This description reveals the essence of the KPCA and KFD methods, but it does not suggest an effective way to implement these two methods, because the direct operation in the high-dimensional or possibly infinite-dimensional feature space is computationally so intensive or even becomes impossible. Fortunately, kernel tricks can be introduced to address this problem. The algorithms of KPCA and KFD can be implemented in the input space by virtue of kernel tricks. An explicit non-linear map and any operation in the feature space are not required at all.

To explain what a kernel trick is and how it works, B. Schölkopf [8] gave an example, as shown in Fig. 1. In the example, the two-class data is linearly non-separable in the two-dimensional input space. That is, one cannot find a projection axis by using any linear dimension reduction technique such that the projected data is separable on this axis. To deal with this problem, the data can be transformed into a feature space by the map Φ given in Fig. 1. As a result, the data become linear separable in the yielding three-dimensional feature space, thereby the linear dimension reduction technique can be applied in such a space. To implement a linear dimension reduction technique in the feature space, one needs to calculate the inner product as follows:

$$\begin{aligned}\langle \Phi(x), \Phi(x') \rangle &= (x_1^2, \sqrt{2}x_1x_2, x_2^2)(x_1'^2, \sqrt{2}x_1'x_2', x_1'^2)^{\mathrm{T}} \\ &= \langle x, x' \rangle^2 \\ &= :k(x, x')\end{aligned}$$

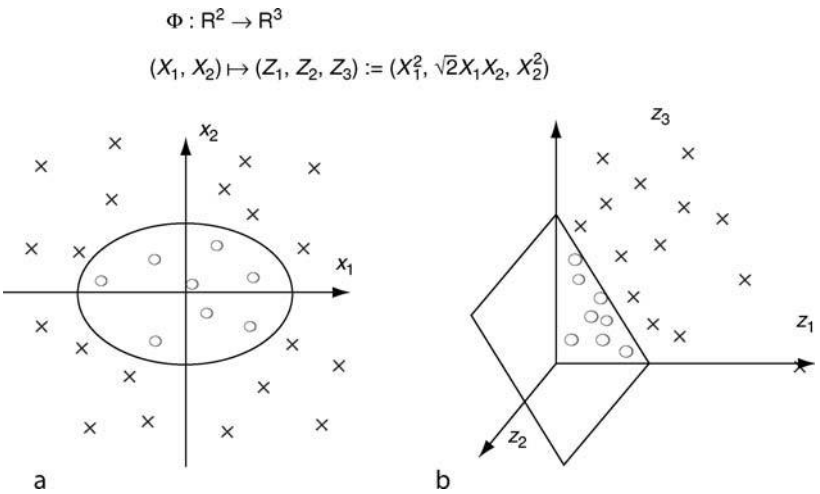
Therefore, the inner product operation can be expressed by a 2-order polynomial kernel function. In this way, the operation of the inner product in the feature space is essentially avoided, as it can be calculated in the input space via a kernel function. In addition, one need not construct an explicit map, since the map is completely determined by the kernel function.

Now, KPCA and KFD can be outlined as follows.

Given a set of M training samples $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$ in $\mathbb{R}^n$, the *covariance operator* on the *feature space* H can be constructed by

$$\mathbf{S}_t^{\Phi} = \sum_{j=1}^{M} \left(\Phi(\mathbf{x}_j) - \mathbf{m}_0^{\Phi}\right)\left(\Phi(\mathbf{x}_j) - \mathbf{m}_0^{\Phi}\right)^{\mathrm{T}} \tag{2}$$

where $\mathbf{m}_0^{\Phi} = \frac{1}{M}\sum_{j=1}^{M} \Phi(\mathbf{x}_j)$, and Φ is a map into the *feature space* H which is determined by a kernel k. In a finite-dimensional Hilbert space, this operator is generally called the covariance matrix.



Non-linear Techniques for Dimension Reduction. **Figure 1** An example of kernel mapping. (**a**) The data is linearly non-separable in the input space. (**b**) The data is linearly separable in the mapped feature space.

It is easy to show that every eigenvector of $\mathbf{S}_t^{\Phi}$, $\boldsymbol{\beta}$, can be linearly expanded by

$$\boldsymbol{\beta} = \sum_{i=1}^{M} a_i \Phi(\mathbf{x}_i). \tag{3}$$

To obtain the expansion coefficients, one can construct the $M \times M$ Gram matrix $\mathbf{K}$ with elements $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ and centralize $\mathbf{K}$ as follows

$$\tilde{\mathbf{K}} = (\mathbf{I} - \mathbf{D})\mathbf{K}(\mathbf{I} - \mathbf{D}), \text{ where } \mathbf{I} \text{ is the identity matrix and } \mathbf{D} = (1/M)_{M\times M}. \tag{4}$$

Calculate the orthonormal eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$ of $\tilde{\mathbf{K}}$ corresponding to the m largest positive eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$. The orthonormal eigenvectors $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ of $\mathbf{S}_t^{\Phi}$ corresponding to the m largest positive eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$ are

$$\beta_j = \frac{1}{\sqrt{\lambda_j}} \mathbf{Q}\,\mathbf{v}_j \; j = 1, \cdots, m$$

where

$$\mathbf{Q} = [\Phi(\mathbf{x}_1), \cdots, \Phi(\mathbf{x}_M)] \tag{5}$$

After the projection of the centered, mapped sample $\Phi(\mathbf{x})$ on to the eigenvector system $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$, one can obtain the KPCA-transformed feature vector $\mathbf{y}$ by

$$\begin{aligned} \mathbf{y} &= (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m)^{\mathrm{T}}[\Phi(\mathbf{x}) - \mathbf{m}_0^{\Phi}] \\ &= \Lambda^{-\frac{1}{2}}\mathbf{V}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}[\Phi(\mathbf{x}) - \mathbf{m}_0^{\Phi}] \\ &= \Lambda^{-\frac{1}{2}}\mathbf{V}^{\mathrm{T}}(\mathbf{I} - \mathbf{D})(\mathbf{K}_{\mathrm{x}} - \mathbf{K}\mathbf{D}_1) \end{aligned} \tag{6}$$

where

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m), \;\; \mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m],$$

and

$$\mathbf{K}_{\mathrm{x}} = [k(\mathbf{x}_1, \mathbf{x}), \;\; k(\mathbf{x}_2, \mathbf{x}), \cdots, k(\mathbf{x}_M, \mathbf{x})]^{\mathrm{T}}$$

and

$$\mathbf{D}_1 = (1/M)_{M\times 1}.$$

KFD seeks a set of optimal discriminant vectors by maximizing the Fisher criterion in the feature space. KFD can be derived in the similar way as used in KPCA. That is, the Fisher discriminant vector can be expanded using Eq. (3) and then the problem is formulated in a space spanned by all mapped training samples. (For more details, please refer to [2, 3].) Recent works [9, 10] revealed that KFD is equivalent to KPCA plus LDA. Based on this result, a more transparent KFD algorithm has been proposed. That is, KPCA is first performed and then LDA is used for a second dimension reduction in the KPCA-transformed space.

Manifold Learning Related Non-Linear Dimension Reduction Techniques

Of late, manifold learning has become very popular in machine learning and pattern recognition areas. Assume that the data lie on a low-dimensional manifold. The goal of manifold learning is to find a low-dimensional representation of data, and to recover the structure of data in an intrinsically low-dimensional space. To gain more insight into the concept of manifold learning, one can begin with a fundamental problem: how is the observation data on a manifold generated? Suppose the data is generated by the following model:

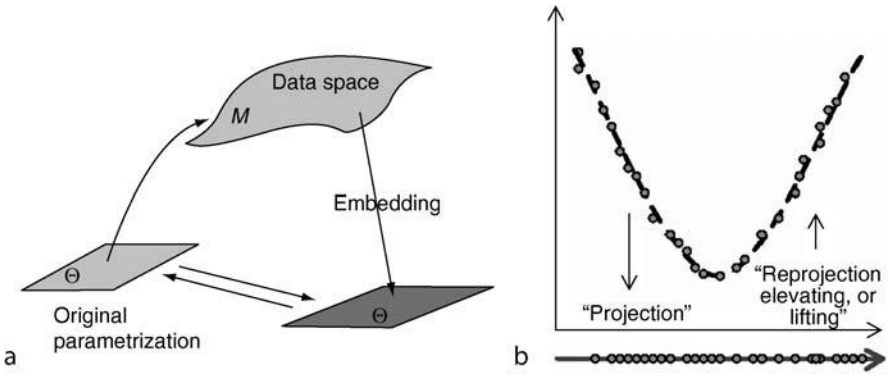
$$f(\Theta) \longrightarrow \mathrm{X}, \tag{7}$$

where Θ is the parameter set, and f can be viewed as a non-linear map. Then, its inverse problem is: how can one recover the parameter set without knowing the map f? How can one build a map f from the data space to the parameter space? Manifold learning seeks answers to these problems. The process of manifold learning is illustrated in Fig. 2; Fig. 2 (a) presents a general process of manifold learning [11], and Fig. 2(b) shows an example of how to find a one-dimensional embedding (parameter set) from the two-dimensional data lying on a one-dimensional manifold [12].

Many manifold learning related dimension reduction techniques have been developed over the last few years. Among the most well-known are isometric feature mapping (ISOMAP) [4], local linear embedding (LLE) [5], and Laplacian Eigenmap [6]. Here, LLE can be considered as an example to introduce manifold learning related dimension reduction techniques.

Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N\}$ be a set of N points in a high-dimensional observation space $\mathbb{R}^n$. The data points are assumed to lie on or close to a low-dimensional manifold. LLE seeks to find a low-dimensional embedding of X by mapping the data into a single global coordinate system in $\mathbb{R}^{\mathrm{d}}$ $(d < n)$. The corresponding set of N points in the embedding space $\mathbb{R}^{\mathrm{d}}$ can be denoted as $Y = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_N\}$.

The LLE algorithm is outlined in the following three steps.



Non-linear Techniques for Dimension Reduction. **Figure 2** Illustration of the process of manifold learning. (**a**) A general process of manifold learning [11]. (**b**) An example of how to derive a one-dimensional embedding from the two-dimensional data lying on a one-dimensional manifold [12].

Step 1: For each data point $\mathbf{x}_i \in \mathrm{X}$, find K nearest neighbors of $\mathbf{x}_i$. Let $\Omega_i = \{j|\mathbf{x}_j$ belongs to the set of K nearest neighbors of $\mathbf{x}_i\}$

Step 2: Reconstruct $\mathbf{x}_i$ from its K nearest neighbors identified. The reconstruction weights can be obtained by minimizing the reconstruction error

$$\varepsilon_i = \left\| \mathbf{x}_i - \sum_j w_{ij}\mathbf{x}_j \right\|^2, \qquad (8)$$

subject to $\sum_j w_{ij} = 1$ and $w_{ij} = 0$ for any $j \notin \Omega_i$.

Suppose the optimal reconstruction weights are $w^*_{ij}(i, j = 1, 2, \cdots, N)$.

Step 3: Compute d-dimensional coordinates $\mathbf{y}_i$ by minimizing the embedding cost function

$$\Phi(\mathbf{y}) = \sum_i \left\| \mathbf{y}_i - \sum_j w^*_{ij}\mathbf{y}_j \right\|^2. \qquad (9)$$

subject to the following two constraints

$$\sum_i \mathbf{y}_i = \mathbf{0} \text{ and } \frac{1}{N}\sum_i \mathbf{y}_i\mathbf{y}_i^T = \mathbf{I}, \qquad (10)$$

where $\mathbf{I}$ is an identity matrix.

The process of LEE algorithm is illustrated in Fig. 1 of SVM and Kernal Method by Schölkopf [8].

Although some previous experiments have demonstrated that the LEE and ISOMAP methods can provide perceptually meaningful representation for facial expression or pose variations, these manifold learning methods may not be suitable for biometric recognition tasks. [13] First, the goal of these manifold learning algorithms has no direct connections to classification. Second, these algorithms are inconvenient to deal with new samples because the involved non-linear map is unknown. How to model biometric manifolds and develop effective manifold learning algorithms for classification purposes deserve further investigation.

Summary

Two kinds of non-linear dimension reduction techniques, kernel-based methods and manifold learning related algorithms, have been introduced here. The mechanism of kernel methods is to increase the dimension first by an implicit non-linear map determined by a kernel and then to reduce the dimension in the feature space, while that of manifold learning related algorithms is to reduce the dimension directly via a non-linear map. Recent research on these two kinds of dimension reduction techniques has revealed an interesting result: manifold learning algorithms, such as ISOMAP, local linear LLE, and Laplacian Eigenmap, can be described from a kernel point of view [14].

Related Entries

- ▶ Biometrics, Overview
- ▶ Kernel Methods
- ▶ Linear Techniques for Dimension Reduction
- ▶ Manifold Learning
- ▶ Non-Linear Dimension Reduction Methods

References

1. Schölkopf, B., Smola, A., Muller, K.R.: Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput. **10**(5), 1299–1319 (1998)
2. Mika, S., Rätsch, G., Weston, J., Schölkopf, B., Müller, K.R.: Fisher discriminant analysis with kernels. IEEE International Workshop on Neural Networks for Signal Processing IX, Madison (USA), August 1999, pp. 41–48
3. Baudat, G., Anouar, F.: Generalized discriminant analysis using a kernel approach. Neural Comput. **12**(10), 2385–2404 (2000)
4. Tenenbaum, J.B., de Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**, 2319–2323 (2000)
5. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science. **290**, 2323–2326 (2000)
6. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. **15**(6), 1373–1396 (2003)
7. Yan, S., Xu, D., Zhang, B., Zhang, H.J., Yang, Q., Lin, S.: Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Trans. Pattern Anal. Mach. Intell. **29**(1), 40–51 (2007)
8. Schölkopf, B.: SVM and Kernel Method. http://www.kernel-machines.org/
9. Yang, J., Jin, Z., Yang, J.Y., Zhang, D., Frangi, A.F.: Essence of Kernel Fisher Discriminant: KPCA plus LDA, Pattern Recogn. **37**(10), 2097–2100 (2004)
10. Yang, J., Frangi, A.F., Yang, J.Y., Zhang, D., Zhong, J.: KPCA plus LDA: A complete kernel fisher discriminant framework for feature extraction and recognition. IEEE Trans. Pattern. Anal. Mach. Intell. **27**(2), 230–244 (2005)
11. Grimes, C., Donoho, D.: Can these things really work? Theoretical Results for ISOMAP and LLE, a presentation at the Workshop of Spectral Methods in Dimensionality Reduction, Clustering, and Classification in NIPS 2002. http://www.cse.msu.edu/∼lawhiu/manifold/
12. McMillan, L.: Dimensionality reduction Part 2: Nonlinear methods. http://www.cs.unc.edu/Courses/comp290-90-f03/
13. Yang, J., Zhang, D., Yang, J.Y., Niu, B.: Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics. IEEE Trans Pattern Anal. Mach. Intell. **29**(4), 650–664 (2007)
14. Ham, J., Lee, D., Mika, S., Schölkopf, B.: A kernel view of the dimensionality reduction of manifolds. In: Proceedings of the Twenty-First International Conference on Machine Learning, Alberta, Canada, pp. 369–376 (2004)

Normalised Hamming Distance

► Score Normalization Rules in Iris Recognition

Nuisance Attribute Projection

► Session Effects on Speaker Modeling

Numerical Standard

Minimal number of corresponding minutiae between a fingermark and a fingerprint necessary for a formal identification, in absence of significant difference.

► Fingerprint, Forensic Evidence of

O

Object Recognition

Given a few training image of the same target object (same object, but may be viewed from different angles or position), the goal of object recognition is to retrieve the same object in other unseen images. It is a difficult problem because the target object in unseen image may appear different from what it appears in the training image, due to the variation of view points, background clutter, ambient illumination, partially occluded by other object or deformation of the object itself. A good object recognition algorithm is supposed to be able to recognize target object given all of the above variations.

► Iris Super-Resolution

Observations from Speech

► Speaker Features

Ocular Biometrics

► Retina Recognition

Odor Biometrics

Adee A. Schoon[1], Allison M. Curran[2], Kenneth G. Furton[2]
[1]Animal Behavior Group, Leiden University, Leiden, The Netherlands
[2]Department of Chemistry and Biochemistry, International Forensic Research Institute, Florida International University, Miami, FL, USA

Synonyms

Osmology; Scent identification line-ups

Definition

Human odor can be differentiated among individuals and can therefore be seen as a biometric that can be used to identify this person. Dogs have been trained to identify objects held by a specific person for forensic purposes from the beginning of the twentieth century. Advancing technology has made it possible to identify humans based on ► headspace analysis of objects they have handled, opening the route to the use of odor as a biometric.

Introduction

From the early twentieth century, dogs have been used to find and identify humans based on their odor. This has originated from the capacity of dogs to follow the track of a person, either by following the odor the person left directly on the ground that the dog needed to follow quite closely ("tracking"), or by following a broader odor trail that the dogs could follow at some distance ("trailing"). Some dogs were very

© 2009 Springer Science+Business Media, LLC

"track-sure": i.e., they continued to follow the specific person in spite of changes in direction, ground surface, and obstacles, in spite of other people having crossed the path earlier or later. Such dogs could also identify the person that had laid that track. This setup is still followed today in the basic training of bloodhounds all over the world. However, a more formalized manner of working with dogs identifying human odors has also evolved, primarily in Europe.

This formalized methodology is called "scent identification line-up," or "osmology", and is applied as a forensic identification tool in several European countries. Dogs are trained to match the odor of a sample to its counterpart in an array of odors. This can be done in different ways [1, 2]. Generally the dog is given a scent sample from a crime scene that presumably contains the odor of the perpetrator. The odor of the suspect and a number of foils, collected in a standardized manner, are offered to the dog as the array. The dog has to match the crime-scene related odor to that of the suspect in the array, and indicate its choice with a learned response. The methods and materials used to collect human odor differ between countries; the exact protocol for working with the dog differs; quality control measures necessary to validate the correctness of the outcome differ; and the way in which the results are evaluated and used during investigation and trial differ between countries too. In spite of efforts to harmonize these differences, they still exist since there is little scientific evidence to select the "best" way: dogs perform best when tested in the way they were trained, and much depends on how the dogs were selected and trained.

From the little scientific work done using dogs in this field, it became clear that dogs are capable of matching odors collected from different body parts [3, 4]. The series of experiments conducted by Schoon and de Bruin [3], showed that trained police dogs were capable of matching objects (stainless steel tubes) held in the pocket or in the crook of the arm to objects held by hand and vice versa significantly better than chance, but that their performance was a lot better on the comparison they trained often (pocket to hand: 58% correct in a 1 out of 6 comparison) than on a comparison they never trained (crook elbow to hand; hand to crook elbow: 32% correct in a 1 out of 6 comparison). Settle [4] had people scenting objects (pieces of gauze) on numerous body parts and also found dogs could match those that had been handled by the same person significantly better than chance (60% correct in a 1 out of 6 comparison). However, the gauzes they used were stored together per person in a glass jar prior the experiments with the dog, so they may have all reached an equilibrium in this jar. Hepper [5] found that dogs use odor cues that are under genetic control more than those under environmental control. He let dogs match the odor of T shirts of fraternal and identical twins with identical or different diets. When both diet and genes were identical, the dogs could not differentiate between the twins (1 out of 2 comparisons). When the genes were identical but the diets differed, the dogs were able to differentiate between the twins but they took a long time and their choices were not very sure (83.5% correct in a 1 out of 2 comparison). When the genes were different but the diets identical, the dogs performed best and made their choices quickly and surely (89% correct in a 1 out of 2 comparison).

With advancing technology in the second half of the twentieth century, an effort was made to identify the source and composition of the body secretions that made it possible for dogs to actually identify people based on their odor. The human skin can be divided into two layers: the outer layer called the epidermis and the inner layer called the dermis. The dermis layer contains most of the specialized excretory and secretory glands. The dermis layer of the skin contains up to 5 million secretory glands including eccrine, apocrine and sebaceous glands [6]. Bacterial breakdown of apocrine secretions result in a huge number of volatile compounds in armpits [7–9], but for forensic purposes the breakdown of sebaceous gland secretions is more interesting since these products can be found on crime-related objects such as guns, knives, crowbars, gloves etc. Further study showed that trained dogs are capable of matching objects scented by the same person at different times but that their performance was lower [10].

Instrumental Differentiation Body Scent

The individual body odors of humans are determined by several factors that are either stable over time (genetic factors) or vary with environmental or internal conditions. The authors have developed distinguishing terminology for these factors: the "primary odor" of an

Odor Biometrics. **Figure 1** Dog searching for a matching odor in a Dutch scent identification line-up (photo courtesy of the Netherlands National Police Agency).

individual contains constituents that come from within and are stable over time regardless of diet or environmental factors; the "secondary odor" contains constituents which also come from within and are present due to diet and environmental factors; and the "tertiary odor" contains constituents which are present because they were applied from the outside (i.e., lotions, soaps, perfumes, etc.) [9]. There is a limited understanding of how the body produces the volatile organic compounds present in human scent. Although the composition of human secretions and fingerprint residues have been evaluated for their chemical composition [6, 7], comparatively little work has been done to determine the volatile organic compounds present in human scent. Knowing the contents of human sweat may not accurately represent the nature of what volatile compounds are present in the headspace above such samples which constitute the scent.

With the use of gas chromatography-mass spectrometry, an increasing number of volatiles were identified in the headspace of objects handled by people [11]. Investigations into the compounds emitted by humans that attract the Yellow Fever mosquito have provided insight into the compounds present in human odor. Samples were collected using glass beads that were rolled between fingers. The beads were then loaded into a GC and cryofocused by liquid nitrogen at the head of the column before analysis with ▶ GC/MS. The results showed more than 300 observable compounds as components of human skin emanations, including: acids, alcohols, aldehydes, and alkanes. The results also showed qualitative similarities in compounds between the individuals studied, however, quantitative differences were also noted [11].

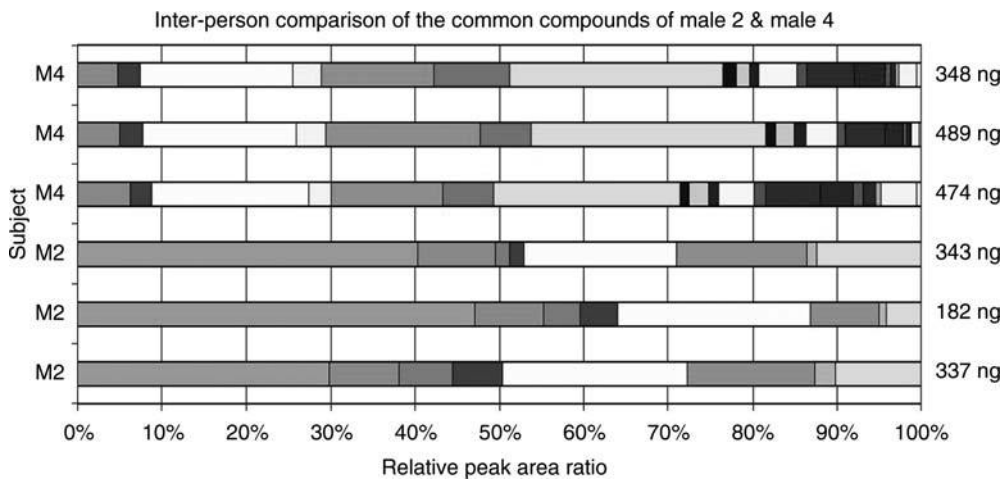
Until recently, technological limitations have restricted the ability of researchers to identify the chemical components that comprise human scent without altering the sample or to use the information to chemically distinguish between individuals. In addition, it has been difficult to distinguish between primary, secondary, and tertiary odor components in a collected human scent sample. ▶ Solid phase micro-extraction (SPME) is a simple solvent-free headspace extraction technique which allows for ▶ volatile organic compounds (VOCs) present in the headspace (gas phase above an item) to be sampled at room temperature. SPME in conjunction with GC/MS has been demonstrated to be a viable route to extract and analyze the VOCs present in the headspace of collected human secretions. In a recent study, the hand odor of 60 subjects were studied (30 males and 30 females) and 63 human compounds extracted, there was a high degree of variability observed with six high frequency compounds, seven medium frequency compounds, and 50 low frequency compounds among the population. The different types of compounds determined to be present in a human hand odor profile included acids, alcohols, aldehydes, alkanes, esters, ketones,

and nitrogen containing compounds. It has been demonstrated that nonparametric methods of correlation can be employed to differentiate between VOC patterns from different individuals. In the 60 subject study, it was shown that Spearman Rank Correlation coefficient comparisons of human odor compounds among individuals is a viable method of data handling for the instrumental evaluation of the volatile organic compounds present in collected human scent samples, and that a high degree of distinction is possible among the population studied [12]. Using a match/no-match threshold of 0.9 produces a distinguished ability of 99.7% across the population. Other work also showed that multiple samples taken from the same person showed that these could not be distinguished at the same level. Figure 2 illustrates the variation of the VOC patterns in multiple samples from two different males.

The genetic source of these specific human volatiles has also been investigated. Experimental work with dogs had already indicated a link to the genes of a person, and work with rats and mice had located the genes of the Major Histocompatibility Complex (MHC) as the source of variation. The genetic basis for individualizing body odors has been studied extensively in genetically engineered mice which differ in respect to the genes present in the MHC [13]. MHC exhibits a remarkable genetic diversity with resulting from a variety of characteristics including a level of heterozygosity approaching 100% in natural populations of mice. This high level of heterozygosity seems to be maintained by behavioral factors including mating success and associated with olfactory cues, and chemosensory imprinting. In humans, the MHC is referred to as the HLA, which is a short for human leucocyte antigen. Experiments utilizing trained rats have shown that urine odors of defined HLA-homozygous groups of humans can be distinguished [13]. Individual body scents of mice can be altered by modification of genes within the MHC. Alterations to the individual body scents of mice result in changes in the concentrations of the volatile components found in the urine [14]. Using two-dimensional GC/MS Willse et al. were able to detect differences in the several dozen MHC compounds (including 2,5-dimethylpyrazine and 2-*sec*-butyl-4,5-dihydrothiazole) found in ether-extracted urine from two inbred groups of mice that differed only in MHC genes.

Legal Perspectives on Human Odor for Forensic Purposes

In Europe, scent identification lineups have been used routinely by police forces, for example in Poland and The Netherlands, and the results have been the subject of discussion and different interpretations in court. In Poland Wójcikiewicz [15] summarized a number of court cases where dog evidence was critically reviewed. Generally, the evidence was accepted by Polish courts as "additional evidence," thus allowing the results to be used only if convergent with other evidence; a point



Odor Biometrics. Figure 2 Illustration of the variety in volatile organic compounds as collected by SPME and determined by GC-MS from three samples of two human subjects. Each color is a different VOC.

of view of Wójcikiewicz, given the limited scientific background knowledge at that time. In the Netherlands, scent lineup evidence has been the subject of much debate over the years. A recent case confirmed that results from carefully conducted scent identification lineups can be used as an addition to other evidence [16]. In the absence of the other evidence, a positive result of such a lineup is regarded as insufficient evidence for conviction.

The twenty-first century has brought with it two important case decisions in the United States Court System pertaining to the use of human scent canines in criminal prosecutions. In 2002, the U.S. Court System decided human scent canine associations could be utilized through the introduction of expert witness testimony at trial if the canine teams were shown to be reliable [17]. In 2005, a Kelley hearing in the state of California [18] set a new precedent in the U.S. which allowed human scent identification by canine to be admitted as forensic evidence in court as opposed to being presented as expert witness testimony. The California court ruled that human scent discrimination by canine can be admitted into court as evidence if the person utilizing the technique used the correct scientific procedures, the training and expertise of the dog-handler team is proven to be proficient, and the methods used by the dog handler are reliable.

Summary

The scientific studies to date support the theory that there is sufficient variability in human odor between persons and reproducibility of primary odor compounds from individuals that human odor is a viable biometric that can be used to identify persons. The bulk of the available literature is based on the ability of training dogs to identify objects held by a specific person but advancing technology has recently made it possible to differentiate humans based on headspace analysis of objects they have handled supporting the results seen with dogs. With additional research and development on training and testing protocols with the dogs, and instrumental methods, the future of human odor as an expanded biometric is quite promising. In addition, unlike many other biometrics, human scent can be detected from traces, such as skin rafts, left by a person and can be collected in a noninvasive fashion.

Related Entries

- ▶ Human Scent and Tracking
- ▶ Individuality

References

1. Schoon, A., Haak, R.: K-9 suspect discrimination: Training and practicing scent identification line-ups. Detselig, Calgery, AB, Canada (2002)
2. Schoon, G.A.A.: Scent identification line-up by dogs (Canis Familiaris): Experimental design and forensic application. Appl. Anim. Behav. Sci. **49**, 257–267 (1996)
3. Schoon, G.A.A., De Bruin, J.C.: The Ability of dogs to recognize and cross-match human odours. Forensic Sci. Int. **69**, 111–118 (1994)
4. Settle R.H., Sommerville, B.A., McCormick, J., Broom, D.M.: Human scent matching using specially trained dogs. Anim. Behav. **48**(6), 1443–1448 (1994)
5. Hepper, P.G.: The discrimination of human body odour by the dog. Perception **17**(4), 549–554 (1998)
6. Ramotowski, R.S.: Comparison of latent print residue. In: Lee, H.C., Gaensslen, R.E. (eds.) Advances in Fingerprint Technology, 2nd edn. pp. 63–104. CRC Press, Boca Raton, FL (2001)
7. Shelley, W.B., Hurley, H.J. Jr., Nichols, A.C.: Axillary odor: Experimental study of the role of bacteria, apocrine sweat and deodorants. AMA Arch. Derm. Syphilol. **68**(4), 430–446 (1953)
8. Sommerville, B.A., Settle, R.H., Darling, F.M., Broom, D.M.: The use of trained dogs to discriminate human scent. Anim. Behav. **46**, 189–190 (1993)
9. Curran, A.M., Rabin, S.I., Prada, P.A., Furton, K.G.: Comparison of the volatile organic compounds present in human odor using SPME-GC/MS. J. Chem. Ecol. **31**(7), 1607–1619 (2005)
10. Schoon, G.A.A.: The effect of the ageing of crime scene objects on the results of scent identification line-ups using trained dogs. Forensic Sci. Int. **147**, 43–47 (2005)
11. Bernier, U.R., Booth, M.M., Yost, R.A.: Analysis of human skin emanations by gas chromatography/mass spectrometry. 1. Thermal desorption of attractants for the yellow fever mosquito (*Aedes aegypti*) from handled glass beads, Anal. Chem. **71**(1), 1–7 (1999)
12. Curran, A.M., Ramirez, C.R., Schoon, A.A., Furton, K.G.: The frequency of occurrence and discriminatory power of compounds found in human scent across a population determined by SPME-GC/MS. J. Chromatogr. **B 846**, 86–97 (2007)
13. Eggert, F., Luszyk, D., Haberkorn, K., Wobst, B., Vostrowsky, O., Eckhard Westphal, E., Bestmann, H.J., Müller-Ruchholtz, W., Ferstl, R.: The major histocompatibility complex and the chemosensory signalling of individuality in humans. Genetica **104**, 265–273 (1999)
14. Willse, A., Belcher, A.M., Preti, G., Wahl, J.H., Thresher, M., Yang, P., Yamazaki, K., Beauchamp, G.K.: Identification of major histocompatibility complex-regulated body odorants by

statistical analysis of a comparative gas chromatography/mass spectrometry experiment. Anal. Chem. **77**, 2348–2361 (2005)
15. Wójcikiewicz, J.: Dog scent lineup as scientific Evidence. Problems of Forensic Sciences **41**, 141–149 (2000)
16. LJN: AW0980, 11th April 2006, Helderse Taximoord
17. California v. Ryan Willis, MA020235, June (2002)
18. California v. Salcido, Cal. App. 2nd, GA052057 (2005)

Off-Angle or Nonorthogonal Segmentation

► Segmentation of Off-Axis Iris Images

On-Card Matching

Chen Tai Pang[1], Yau Wei Yun[1], Xudong Jiang[2]
[1]Institute for Infocomm Research, A*STAR, 21 Heng Mui Keng Terrace, Singapore
[2]Nanyang Technological University, 50 Nanyang Avenue, Block S2-B1c-105, Singapore

Synonyms

Biometric Match-on-Card, MOC; Work-Sharing On-card Matching

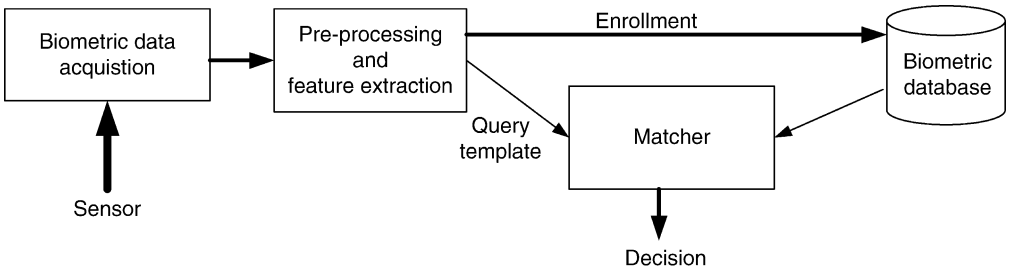
Definition

On-card matching is the process of performing comparison and decision making on an integrated circuit (IC) card or smartcard where the biometric reference data is retained on-card to enhance security and privacy. To perform enrolment, the biometric interface device captures the biometric presentation of the user to create the biometric ► template. Then, the biometric template and user's information are uploaded to the card's secure storage. To perform on-card matching, the biometric interface device captures the biometric presentation and creates a biometric template. The created biometric template is then uploaded to the card for verification. The verification process shall be executed on-card instead of sending the enrolled template out of the card for verification.

Introduction

The need for enhanced security persists more than ever in a more electronically dependent and interconnected world. The traditional authentication method, such as PIN, is neither secure enough nor convenient for automatic identification system such as border control. Our economic and social activities in today's electronic age are getting more reliant to electronic transactions that transcend geological and physical boundaries. These activities are supported by implicitly trusting the claimed identity – with we trusting that the party we are transacting with is genuine and vice versa. However, conventional password and Personal Identification Number (PIN) commonly used are insecure, requiring the user to change the password or PIN regularly. Biometric technology uses a person's unique and permanent physical or behavioral characteristics to authenticate the identity of a person. Higher level of security can be provided for identity authentication than merely the commonly used PIN, password or token. Some of the popular biometric technologies include fingerprint, face, voice, and iris. All biometric technologies share a common process flow as shown in (Fig. 1) below.

Fig. 1 shows the basic architecture of biometric authentication with a central database. In order to use the biometric system to identify a person, he or she will have to enroll in the system's database. The system has to create and maintain the biometric database in a central PC or server. Even for a biometric door access system (no matter for home use or office use), a small biometric database is stored in the embedded unit. Usually this is not a problem for home use because only the owner or trusted person can have access to the database. But what about the other service providers? If hackers can access some of the confidential database information of big corporations such as Bank of America, LexisNexis, T-Mobile [1] and the security breach affecting more than 200,000 credit card holders [2] who then can the user trust? Since biometric data is permanent and each person has limited amount of choice (a person only has a face and 10 fingers), having the biometric database information stolen is a serious implication to the actual owner. One of the alternatives is to store the biometric template into a smartcard. Smartcard is a plastic card with microprocessor inside to handle the data storage and has processing capability with security features. Hence,



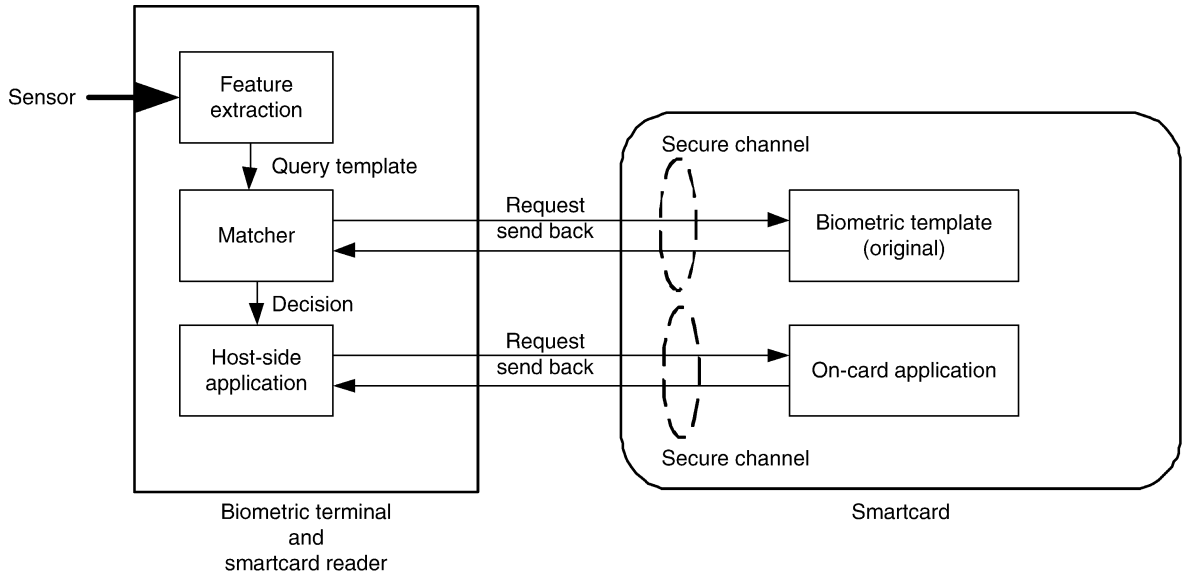
On-Card Matching. Figure 1 Process flow involved in a common biometric system.

the combination of biometrics and smartcard offers enhanced security for identity authentication.

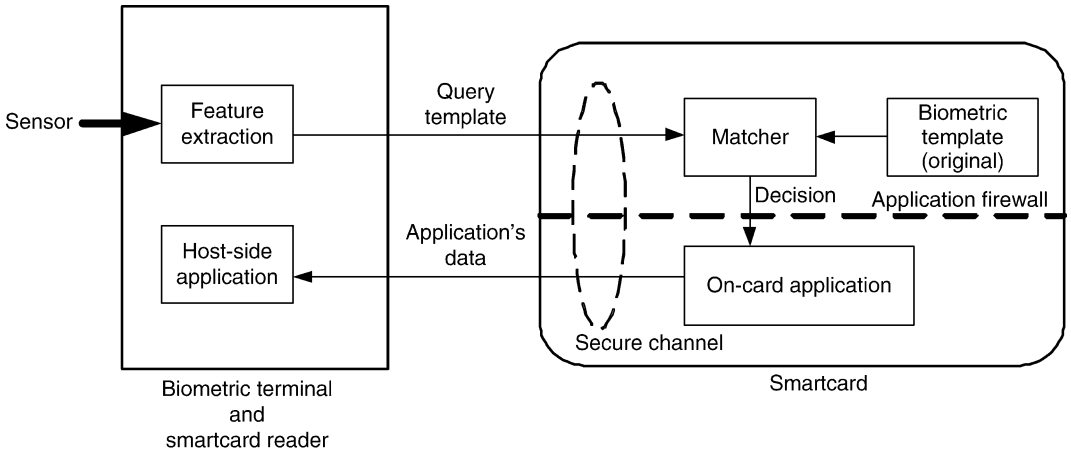
Biometrics and Smartcard

Instead of relying on a centralized database system and allowing individual service provider to create its own biometric database, the biometric information can be kept in the hand of the respective owner of the biometric data. This can be done by putting the biometric data into a secure storage such as a smartcard. Smartcard is a plastic card with an embedded microprocessor, memory and, security features. The user can conveniently carry the smartcard, and thus it also offers mobility to biometric data. The combination of biometric and smartcard offers the advantages of mobility, security and strong identity authentication capability and, at the same time offers the user, a high degree of control over who have access to that biometric data. Hence, biometrics on the smartcard can minimize the privacy concern. There are four distinct approaches to combine the smartcard and biometric technologies as follows:

1. *Template-On-Card (TOC):* This type of matching is also known as off-card matching. The entire process of biometric data acquisition, feature extraction, and matching is done at the terminal or reader side. However, during the enrolment stage, the original template which is constructed at the reader is stored inside the smartcard. During matching, the reader will request for the original template to be released from the smartcard which is then matched with the query template. The decision of further accessing information from the smartcard is made on the reader side. The smartcard itself act as a storage device. Cryptography should be used to mutually authenticate the card and the biometric interface device. To protect the communication between the biometric interface device and the card; a secure channel should be established prior to the transfer of any template or data. As the biometric template and other data objects such as passport/visa or financial account information are stored as a separate file in the smartcard, separate secure channels can be used for transmitting different data object. Fig. 2 shows the basic architecture of TOC.
2. *Match-On-Card (MOC):* MOC means the biometric verification is performed in the card. The process of biometric data acquisition and feature extraction is done at the biometric terminal. During the initial enrolment stage, the original template constructed at the reader is stored inside the smartcard. During matching, the reader will construct the query template which is then sent to the smartcard for matching. The final matching decision is computed inside the smartcard and thus the entire original template is never released from the smartcard. Fig. 3 shows the authentication process of a MOC system for a simple case of border control transaction. The dotted line in the figure is the applet firewall which restricts the access to the matching applet to enquire the status of fingerprint authentication. Therefore, the matching result will be sent from the Matcher to the on-card application by secured sharable method via smartcard operating system. Neither the original template nor the matching result is revealed to the outside world. In order to protect the communication between the biometric interface device and the card, a secure and trusted channel is required.
3. *Work-Sharing On-Card Matching:* ▶ Work-sharing on-card matching is similar to on-card matching except for extra matching procedures are involved



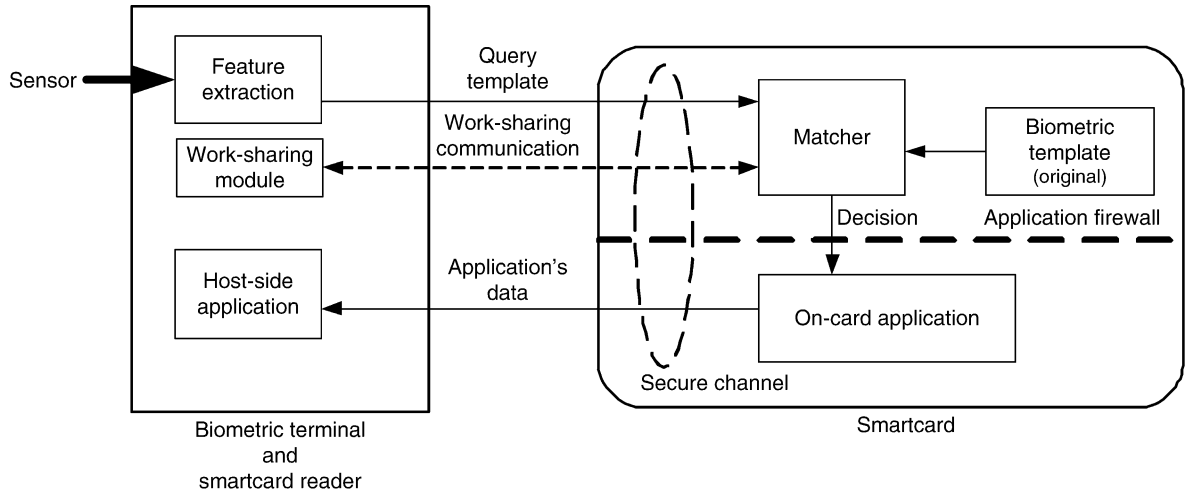
On-Card Matching. **Figure 2** Template-on-card authentication.



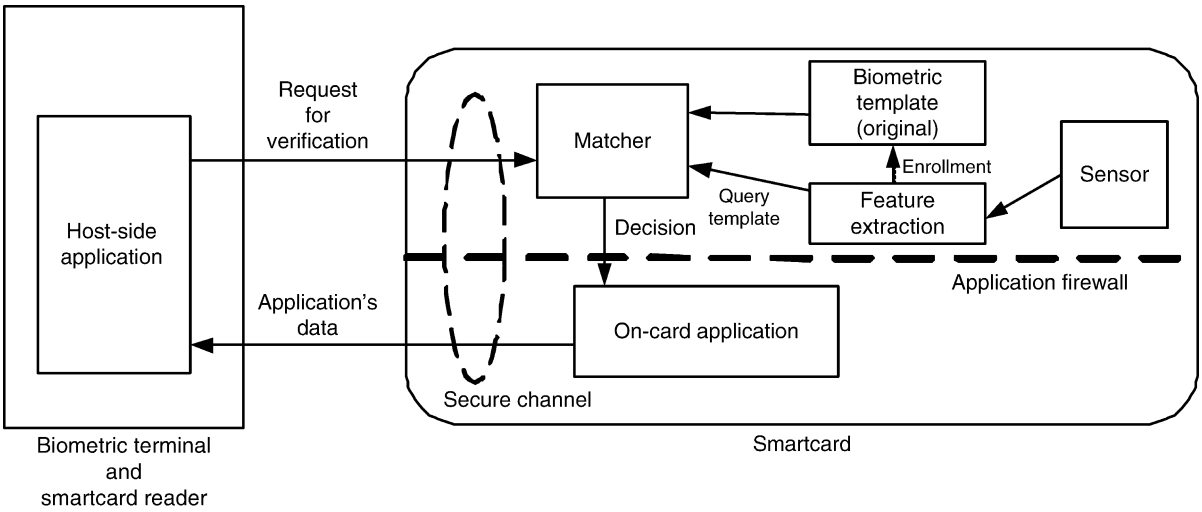
On-Card Matching. **Figure 3** Match-on-card authentication.

to speed up the process. This type of matching is designed for those cards which do not have sufficient processing power and resources to execute the biometric matching. In this case, certain parts which are computation intensive such as template alignment, are sent to the biometric terminal via communication channel to perform computation. The computed intermediate result is sent back to the smartcard to continue with the matching process. The final calculation of the matching score shall be calculated inside the smartcard. Establishing a secure channel is required to protect the communication between the biometric terminal and the smartcard. Fig. 4 shows the basic architecture of work-sharing on-card matching.

4. *System-On-Card (SOC):* ▶ System-on-card matching means the whole biometric verification process, including the acquisition, is performed on the smartcard. The smartcard incorporates the entire biometric sensor, with processor and algorithm. Therefore, the entire process of biometric data acquisition, feature extraction, and matching is done inside the smartcard itself. Both the original template and the query template are computed



On-Card Matching. **Figure 4** Work-sharing On-card matching.



On-Card Matching. **Figure 5** System-on-card authentication.

in the smartcard and do not leave the card. Fig. 5 shows the general authentication process of a SOC system.

Advantages of Match-on-Card

The level of security of a biometric system is judged by examining where the feature extraction and matching takes place. From the point of view of security, system-on-card (SOC) offers the strongest security while template-on-card (TOC) offers the weakest secure for token based authentication [3]. It is obvious that the SOC offers the highest security since the biometric authentication process, including acquisition of biometrics, is executed inside the smartcard itself and no biometric data is transferred out of the smartcard. However, the cost of such smartcard will be high since the card contains a biometric sensor and requires a powerful processor (usually 32-bit) to meet the computational demand of the biometric processing. Therefore, SOC is still not practical for mass issuing and is usually suitable for vertical market only. This means that the match-on-card (MOC) technology which offers a higher security than the TOC technology at reasonable price and is a more practical solution. There are a lot of commercial implementations for fingerprint, face, and iris. Fingerprint MOC is the most popular in market due to good

accuracy, ease of use, affordability, and an overall compact solution.

The reasons why the match-on-card (MOC) technology provides better security in comparison to template-on-card (TOC) technology are:

1. *Better Security and Privacy Protection:* TOC needs to send the enrolment template from the card to the biometric terminal for verification. The security is compromised due to information exposure. Even though the template is usually encrypted, the on-card crypto engine is usually not very strong due to constrained hardware specification of the smartcard's CPU. For the MOC case, the reader will send the query template to the smartcard for identity verification. Therefore, the MOC technology does not reveal the entire original biometric template stored in the smartcard. During the matching process, the stored original template is always trusted since the smartcard is considered a secure storage device. Moreover, better privacy protection can be provided by match-on-card as no one can download the user's enrolment fingerprint template from the card.
2. *Two Factor Authentication:* MOC technology will establish a true two-factor authentication process for the identity authentication needs. No matter MOC or TOC, to start communication between smartcard and reader securely, a secure channel shall be established with mutual authentication before any transaction takes place. This stage is to allow the reader and the smartcard to verify the cryptogram from each side to ensure both reader and smartcard are valid and genuine. However, this stage relies on exchanging challenge code between card and reader. Once the challenge code is stolen by Trojan, hacker may be able to access the smartcard and continue to do further hacking procedures. For TOC, if the first stage is cracked, the hacker will be able to access secured information in the card. For MOC, if the first stage is cracked, the hacker will still need to hack the second stage of biometric MOC stage in order to continue to access secured information. Hence, MOC offers true two-factor authentication which can provide stronger security to protect against hacking.
3. *On-Card Decision Making, Stronger Software Security:* In figure 3, the on-card matcher sends the decision to other on-card application internally via a software firewall that is controlled by the smartcard Operating System (OS). Such internal decision passing via firewall is a strong security feature and very difficult to be hacked. Note that the installation of on-card application is usually done in the factory (ROM masking), in OS provider of the smartcard or in authorized agency with security code for installation. After installing all necessary applications, it is possible to lock the card forever to prevent installing other application in the future. Each application has restriction to access resources from other applications and usually controlled by the smartcard OS. Among the trusted applications, they can send and receive information among them via the firewall with security code. Hence, it is very difficult for hacker to upload Trojan to the card to hack the internal invocation between applications, stealing internal information from the card and sending fake decision from the MOC to fool other on-card applications to leak crucial information.

Implementations of Fingerprint Match-On-Card

In recent years, there are quite a number of attempts to design algorithm to perform fingerprint match-on-card application. Mohamed [4] proposed a memory efficient scheme of using line extraction of fingerprint that could speed up the matching process. However, this approach still needs a 32-bit DSP to process and the computation is still relatively intensive for commercial a smartcard. Vuk Krivec et al. [5] proposed a hybrid fingerprint matcher, which combines minutiae matcher and homogeneity structure matcher, to perform authentication with smartcard the system. Their method is to perform minutiae match-on-card first. Upon successful minutiae matching, the card delivers rotational and translational parameters to the system to perform second stage homogeneity structure fingerprint matcher on the host side. However, this hybrid approach cannot increase the accuracy significantly compared to minutiae matcher alone but using extra time to perform extra host side matching. Andy Surya Rikin et al [6] proposed using minutia ridge shape for fingerprint matching. The ridge shape information is used during the minutiae matching to

improve the matching accuracy. In their experiment, only 64 bytes per template was used. They showed that the accuracy was comparable with the conventional matching but having a faster matching speed. The matching time on a 16-bit smartcard was around 1.2 seconds with 18 minutiae. Mimura M. et al. [7] described a method of designing fingerprint verification on smartcard with encryption functions to enable application using on-card biometrics to perform transaction via Internet. Stefano Bistarelli et al. [8] proposed a matching method using local relative information between nearest minutiae. This method could achieve matching time from 1 to 8 seconds with 10% ERR on average using FVC2002 database. All the above attempts were to implement fingerprint matching on native smartcard or Java card in the research community. Generally speaking, it is not easy to achieve good accuracy with low computation requirement for on-card fingerprint matching. Besides good matching algorithms, software optimization is also an important criterion to develop MOC system to achieve fast on-card matching speed.

Of course, there are several commercial implementations for fingerprint MOC. Most of them are using minutiae data for verification of identity. Those companies usually provide the accuracy information of False Acceptance Rate = 0.01% and False Rejection Rate = 0.1%. No further information regarding the database, method of calculation, and other details have been disclosed. Hence, it is not possible to tell the actual accuracy of those commercial implementations using their provided specification. Currently, the only reliable benchmarking is using common database such as Fingerprint Verification Competition (FVC) fingerprint database or National Institute for Standardization and Technologies (NIST) fingerprint database to compare the other system by using common performance indicators such as False Match Rate (FMR), False Non-Match Rate (FNMR), Equal Error Rate (ERR) and Receiver Operation Curve (ROC) to compare the relative performance among MOC implementations.

Performance of Fingerprint Match-on-card

In 2007, NIST conducted an evaluation for the performance of fingerprint match-on-card algorithms - MINEX II Trial. The aim of MINEX II trial was to evaluate the accuracy and speed of the match-on-card verification algorithms on ISO/IEC 7816 smartcards. The ISO/IEC 19794-2 compact card fingerprint minutiae format was used in the test. The test was conducted in 2 phases. Phase I was a preliminary small scale test with release of report only to the provider. Phase II was a large scale test for performance and interoperability. Initially, 4 teams participated in the Phase I. In the final Phase II test, three teams were participated in the test. The Phase II report was published on 29th February 2008 [9]. Some highlights of the result are stated below:

- The most accurate match-on-card implementation executes 50% of genuine ISO/IEC 7816 VERIFY commands in 0.54 seconds (median) and 99% within 0.86 seconds.
- The False Non-Match Rate (FNMR), at the industrial preferred False Match Rate (FMR) = 0.01%, is 2 to 4 times higher than FMR at 1%.
- Using OR-rule fusion at a fixed operating threshold, the effect of using a second finger only after a rejection of the first, is to reduce false rejection while increasing false acceptance.
- The most accuracy implementation satisfies only the minimum requirements of the United States' Government's Personal Identity Verification (PIV) program.
- Some cards are capable of accepting more than 60 minutiae for matching. Some cards need minutiae removal for either or both of the reference and verification templates prior to transmission to the card. It was discovered that the use of minutiae quality values for removal is superior to using the radial distance alone.

In this evaluation, only 1 team can achieve the minimum requirement of PIV program. Hence, compared to off-card matching, it is necessary to further improve the accuracy for those applications that require PIV specification such as immigration. As the compact card format is the quantized version of the normal size finger minutiae format, the performance is still unknown of using the normal format in MOC. Number of existing commercial implementations are using fingerprint minutiae proprietary format for MOC implementation. MINEX II continues the Phase III in 2008 to gauge improvements over existing implementations and to evaluate others.

Standardization

In order to allow for better global interoperability, several efforts to standardize the biometric match-on-card technology are on-going. There is effort at the international standards body ISO/IEC JTC1 SC17 WG11 to introduce the match-on-card standards. In addition, the effort on biometrics is also on-going, such as at the ISO/IEC JTC1 SC37 level to develop the compact fingerprint template format suitable for smartcards.

In 2005, several standards for biometric data interchange format have been published including finger minutiae data, finger pattern spectral data, finger image data, face image data, iris image data, finger pattern skeletal data etc. ISO/IEC 19794-1 [10] is intended to describe the framework for defining biometric data interchange formats. In ISO/IEC 19794-2 finger minutiae data [11], compact card format is included in the specification to support fingerprint authentication with smartcard. The document ISO/IEC 7816-11 [12], published in 2004, specifies basic operations for performing personal verification through biometric methods using smartcard. However, the above standards are not sufficient for biometric match-on-card. A standard with more in-depth specification is needed for deployment of match-on-card with better interoperability.

In 2006, a new work group (WG) 11 in Sub-Committee 17 under Joint Technical Committee 1 (JTC1) of ISO/IEC was formed. The role of WG11 is to define the functional blocks and components for the use of integrated circuit (IC) cards in applications where the matching of biometric identifiers is to be performed on-card. The document entitled "24787 Information technology - Identification cards: On-Card matching" is still under committee draft stage as of February 2008.

Technical Challenges

MOC technology is challenging to develop due to the limited resources – computational power, memory, and power supply, in the smartcards. For example, today's PC has powerful specification while for smartcard has relatively much lower processing capability. For example, one of the high-end configurations is only 16-bit, 25 MHz processor with 8Kb RAM and 1Mb flash memory. There are few 32-bit smartcards but the price are quite expensive. The most widely used smartcard is the 8-bit card due to its low cost. Moreover, applications in the smartcard have to share resources especially the limited static memory for runtime execution. For a contactless smartcard, availability of RF power is crucial. If the peak power demanded by the intensive computation is not met or that the computational duration is longer than what the power can be sustained from the reader through induction, then the matching process will fail. As a consequence, the user will experience a "false rejection" even though the rejection is not due to the outcome of the biometric matching. Moreover, software optimization is very crucial for MOC implementation to achieve good matching performance. The optimization in term of speed, resource allocation and code size are necessary during the system design phase. Nevertheless, fingerprint match-on-card can already be realized today on an off-the-shelf Java card having 8-bit, 5 MHz CPU core, 5k bytes RAM and 32Kb EEPROM or better with Java OS. The following optimization methods are commonly employed in development of match-on-card technology:

1. Reduce the size of the template: Reduce the amount of information to be matched during on-card matching can reduce the overall matching time. For example, fingerprint match-on-card can restrict the maximum number of minutiae to less than 60 minutiae to be matched per template. However, information reduction may degrade the accuracy of the matcher. Developer should be aware of how much information should be reduced to achieve acceptable accuracy.
2. Work-sharing biometric match-on-card: Some low-end smartcards are not able to handle the whole biometric matching algorithm within acceptable timing. In this case, work-sharing architecture which has been introduced previously can be used to speed up the matching process. The idea of work-sharing architecture is to assist the smartcard to compute those computation intensive functions of the matching algorithm, such as template alignment, using the biometric terminal. The final biometric comparison, such as the calculation of matching score, shall be computed inside the smartcard. The smartcard can send intermediate data or information other than enrolment template to the terminal using secure channel. The developer should be aware of the security

requirement to design matching algorithm using work-sharing architecture for smartcard.

3. Biometric codeword or hashing: Some researchers investigate algorithms to generate codeword such as finger code [13] or biohashing [14]. In this case, the complexity of the matching algorithm in the smartcard can be reduced. However, the stability of the biometric codeword is still not as robust as fingerprint matching due to alignment and deformation of the biometric presentation. Hence, developer should be aware of whether the accuracy is sufficient for particular application.

Summary

Biometric match-on-card technology holds great promise in offering good security and privacy protection. The technology has come a long way to become feasible today at an attractive cost and more can still be done to make it better and cheaper. It provides a good platform for the launch of a nation wide strong identity authentication capability which will open up many other new applications and business possibilities that will provide better convenience, security and protection to the users as compared to what is being used today. There is also a foothold of this technology in the global push for machine readable travel documents which hopefully will lead to a global opportunity in biometric system level application.

Related Entries

- ▶ Authentication
- ▶ Enrolment
- ▶ Identification
- ▶ Match-on-Card
- ▶ Verification

References

1. Hines, M.: LexisNexis break-in spurs more calls for reform. ZDNet.com, http://news.zdnet.com/2100-1009-5606911.html (2005)
2. Vijayan, J.: Scope of credit card security breach expands. Computerworld, http://www.computerworld.com/securitytopics/security/ story/0,10801,101101,00.html (2005)
3. Scheuermann, D.: Identification cards - Integrated circuit cards with contacts - Part 11: Personal verification through biometric methods, ISO/IEC 7816-11:2004. International Organization for Standardization/International Electrotechnical Commission (2006)
4. Allah, M.M.A.: A fast and memory efficient approach for fingerprint authentication system. IEEE Conf. Adv. Video Signal Based Surveill. pp. 259–263 (2005)
5. Krivec, V., Birchhauer, J., Marius, W., Bischof, H.: A hybrid fingerprint matcher in memory constrained environments **2**(18-20), 617–620 (2003)
6. Rikin, A.S., Li, D., Isshiki, T., Kunieda, H.: A Fingerprint Matching Using Minutia Ridge Shape for Low Cost Match-on-Card Systems. IEICE Trans. **E88-A**(5), 1305–1312 (2005)
7. Mimura, M., Ishida, S., Seto, Y.: Fingerprint verification system on smartcard. In: International Conference on Consumer Electronics pp. 182–183 (2002)
8. Bistarelli, S., Santini, F., Vaccarelli, A.: An Asymmetric Fingerprint Matching Algorithm for Java Card$^{TM.}$ Pattern Anal. Appl. J. **9**(4), 359–376 (2006)
9. Grother, P.W.S., Watson, C., Indovina, M., Flanagan, P. (eds.): Performance of Fingerprint Match-on-Card Algorithms Phase II Report. National Institute of Standards and Technology (2008)
10. Stardards: Information technology - Biometric data interchange formats - Part 1: Framework, ISO/IEC 19794-1:2006. International Organization for Standardization/International Electrotechnical Commission (2006)
11. Stardards: Information technology - Biometric data interchange formats - Part 2: Finger minutiae data, ISO/IEC 19794-2:2006. International Organization for Standardization/International Electrotechnical Commission (2006)
12. Stardards: Usability of Biometrics in Relation to Electronic Signatures, ISO/IEC JTC1/SC17 N1793. International Organization for Standardization/International Electrotechnical Commission (2000)
13. Jain, A.K., Prabhakar, S., Hong, L., Pankanti, S.: Filterbank-based Fingerprint Matching, IEEE Transactions on Image Processing. IEEE Trans.Image Process. **9**(5), 846–859 (2000)
14. Teoh, B.J.A., Ngo, C.L.D., Goh, A.: BioHashing: two factor authentication featuring fingerprint data and tokenised random number. Pattern Recognit. **37**, 2245–2255 (2004)

One-to-Many Identification

One-to-many matching is a process of searching a dataset to match the target image with those of more than one persons to identify an individual. See also ▶ Identification.

If a biometric sample is compared with n templates, where n is a positive integer, the matching process is

referred to as a *1:n matching*. In the case of $n = 1$, it is equivalent to a verification process. If $n > 1$, it is regarded as an identification process and sometimes expressly described as 1:many matching.

► Biometrics, Overview

One-to-One Verification

One-to-many matching is a process of matching the target image against those of the claimed person to verify an individual. See also ► Verification.

► Biometrics, Overview

Online Learning

► Incremental Learning

Open-Set Identification

It is unknown whether the subject presented to the biometric system for recognition has enrolled in the system or not. Therefore, the system needs to decide whether to reject or recognize him as one of the enrolled subject. It is the opposite of "Closed-Set Identification."

► Performance Evaluation, Overview

Operational Tests

Operational tests are those in which a biometric system collects and processes data from actual system users in a real field application. Operational tests differ fundamentally from technology and scenario tests in that the experimenter has limited control over data collection and processing. Because operational tests should not interfere with or alter the operational usage being evaluated, it may be difficult to establish ground truth at the subject or sample level. As a result, operational tests may or may not be able to evaluate match or enrollment rates, FRR, or FTE; instead they may be able to directly measure acceptance rates (without distinction between genuine and impostor) and operational throughput.

► Performance Testing Methodology Standardization

Operational Times

Stephen J. Elliott[1], Eric P. Kukula[1], Richard T. Lazarick[2]
[1]Department of Industrial Technology, Purdue University, West Lafayette, IN, USA
[2]Chief Scientist-GSS Identity Labs Computer Science Corporation Camp Hill, PA

Synonyms

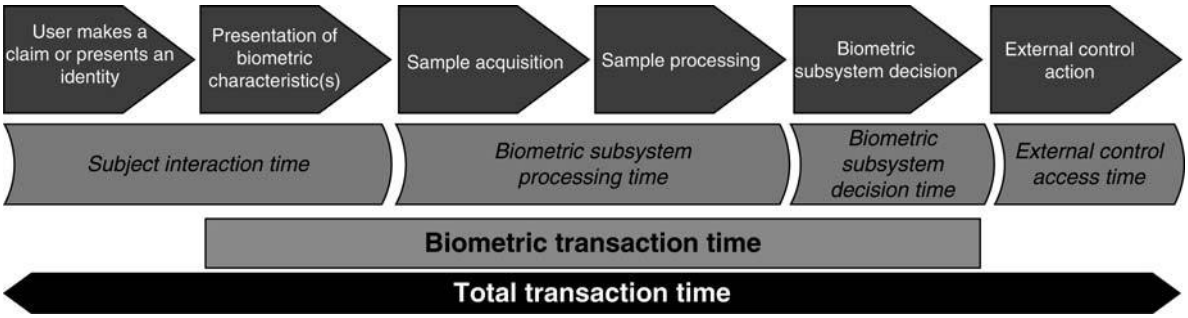
Biometric decision time, and the external operation time; Biometric subsystem transaction time; Biometric transaction time; External operation time; Subject interaction time; Total transaction time

Definition

There are a number of associated definitions for "time" as they relate to an operational biometric system(s). These "time" metrics include the total transaction time, the overt biometric transaction time, the subject interaction time, the biometric subsystem processing/transaction time, the biometric decision time, and the external operation time.

Introduction

One definition of time is given as "the length of time taken to complete an activity" [1]. For a biometric system, "time" can be segmented in alignment with a particular activity or function. Assuming that the biometric model in Fig. 1 is for an access control system,



Operational Times. Figure 1 Types of transaction times.

and that the steps of that model are shown, there are five different types of event time. These include the subject interaction time, the biometric subsystem processing time, the biometric decision time, the biometric transaction time, and the total transaction time.

This document outlines the various timing metrics that are used in biometric verification applications. They include the overall transaction time, the subject interaction time, the biometric subsystem processing time, and the biometric decision time. Although the document includes examples from a technical contribution presented to Working Group 5 in the international biometric standards committee (ISO/IEC JTC 1 SC37) on the developments of timing metrics in scenario testing for biometric access control systems, the metrics can be applied to generic biometric verification systems.

Total Transaction Time

The total transaction time is a sum of all the subcomponent periods of time associated with the biometric application system. For a biometric verification system, the overall transaction time is initiated when the user makes a claim or presents an identity (i.e., swipes a card or enters a PIN). The overall transaction time is completed when the last measurable component has been satisfied, in the physical access control system when the door strike is activated.

Biometric Transaction Time

This begins with the biometric sample presentation and ends with the biometric decision. Therefore, this includes the presentation of the biometric trait portion of the subject interaction time, biometric subsystem processing time, which includes sample acquisition and sample processing time, and the biometric decision time.

Subject Interaction Time

Using the same access control system as described earlier, the subject interaction time commences when a claim of identity is made (or presented), that is, swiping a card or entering a PIN by the user. The time ends when the individual has presented his/her biometric characteristic(s) and the sensor begins to acquire the sample.

Biometric Subsystem Processing Time

The biometric subsystem processing time is the time taken for the system to acquire the biometric sample, to evaluate the quality of the sample, and if the quality is satisfied, to process that sample for comparison. For the samples of bad quality, the biometric system requests the subject to submit the biometric trait. The biometric subsystem processing time ends when either a comparison score or a request for re-submission is generated.

Biometric Decision Time

The biometric decision time is the time required by the biometric subsystem to generate an accept or reject response based on the comparison score and the decision logic. The decision logic could be a simple threshold or a more complex methodology such as fusion logic. In biometric identification where the biometric subsystem generates a list of matched candidates, the biometric decision time is the time required to search

the subject from the database of enrolled subjects. In this case, it depends on both the size of the database and the strategy in search, for example, the classification or binning of biometric samples.

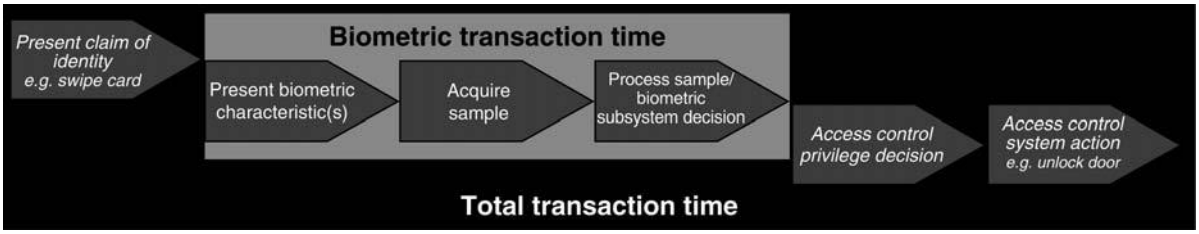
External Operation Time

Provided that the biometric decision is an accept, the external operation time is the time required to complete the application transaction. In the physical access control system, it is the time required by the physical electro-mechanical components to act according to the decision of the biometric subsystem and other access control privilege criteria to complete the access transaction. In biometric identification, the external operation time is not required because the external operation is typically a manual function.

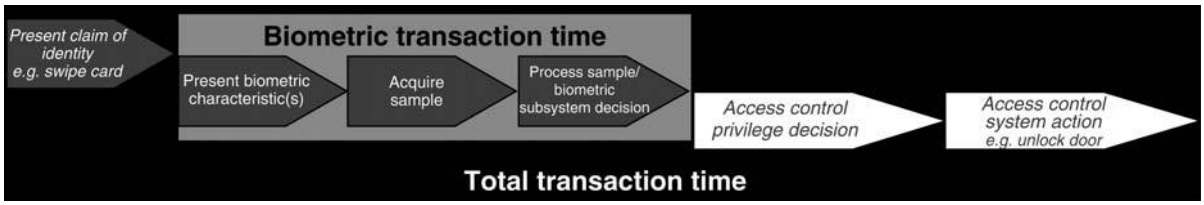
Timing Illustrations

According to a Technical Contribution from the ISO/IEC JTC 1 SC 37 WG5 Special Group on 19795–5 [2], timing can be related to the type of testing to be undertaken. There are three major types of testing: technology, scenario, and operational.

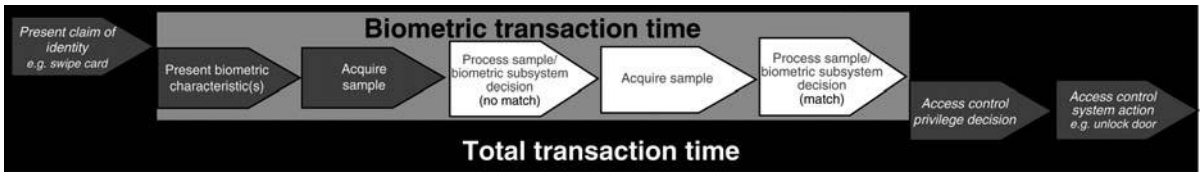
For the purpose of evaluating the performance of a biometric subsystem in a scenario evaluation, it is appropriate to define a metric that only measures the biometric subsystem performance and is not influenced by the presence or performance of other access control devices, policies, and functions. Thus, the distinction is made between biometric transaction time and total transaction time. This distinction is made in timing diagrams illustrating the components of a biometric transaction and a total access control system transaction shown in Figs. 2–4.



Operational Times. **Figure 2** Timeline showing total and biometric subsystem transaction times for a one-attempt transaction.



Operational Times. **Figure 3** Timeline showing total and biometric subsystem transaction times for a one-attempt transaction with longer access control time.



Operational Times. **Figure 4** Timeline showing total and biometric subsystem transaction times for a two-attempt transaction.

Related Entries

- ▶ Attempt
- ▶ Failure to Acquire (FTA)
- ▶ Failure to Enroll (FTE)
- ▶ General Biometric Model
- ▶ Match rates
- ▶ Performance Testing and Evaluation – Technology, Scenario, and Operational Testing
- ▶ Presentation
- ▶ Threshold
- ▶ Transaction

References

1. Soanes, C., Hawker S.: Compact Oxford English Dictionary of Current English (3rd ed.). Oxford University Press, New York (2005)
2. International Organization for Standardization.: *ISO/IEC JTC1 SC37* Comments from the Special Group Established to Review 19795–5 (N-1770) (No. SC37N1853). Geneva: ISO/IEC (2006)

Optical Flow

Optical flow is a vector field that represents the motion of pixels between two images. A variety of algorithms have been developed for computing optical flow. It is commonly computed for video, in which the motion from one frame to the next is relatively small.

- ▶ Face Variation
- ▶ Gait Recognition, Silhouette-Based

Optical Target

Optical target refers to an object the optical characteristics of which are well known. It is used to perform an optical calibration of lenses or illuminators.

- ▶ Biometric Sensor and Device, Overview

Optimal Hyperplane

- ▶ Support Vector Machine

Optimization

Given a cost function, different strategies could be used to obtain the estimate. This is called the optimization strategy and the solution often depends upon the exact strategy that is used.

- ▶ Face Tracking

Ordinal Measure

In ordinal measure, variables are required to be monotonically related so that they can be rank-ordered. In palmprint identification, ordinal measure qualitatively compares neighborhood image pixels or regions and preserves their ordinal relationship. This yields a symbolic representation of the relations.

- ▶ Palmprint Features

Orthographic Scanning

The procedure of recording two-dimensional images of an object with the goal of capturing the object's three-dimensional structure. In a hand-geometry device, for example, it is performed with the help of a mirror that projects the lateral surface of the hand into the visual field of the camera and allows it to record both the side and top views of the hand in a single image.

- ▶ Hand-Geometry Device

Osmology

Osmology is the study of odors and the sense of smell; phrase coined by the Polish Forensic Police for the field of science where trained dogs compare scent traces that the perpetrator of a crime leaves at the crime scene to the odor of a person suspected of that crime in a line-up procedure.

► Odor Biometrics

Output Noise Variance (ONV)

The output noise variance is the variance of the noise of a correlation plane. If the input noise is stationary and with Gaussian distribution, the output noise on correlation plane is also going to be stationary and Gaussian distributed. The output noise variance describes the output noise variance at all pixels in the output. If we would like to see a noise-tolerant peak on the correlation plane for the authentic comparison, it is desirable to make the noise variance on the correlation plane as small as possible on average. That is why would like to incorporate the criteria for minimizing ONV in the design of the correlation filters.

► Iris Recognition Using Correlation Filters

Outsole Pattern Matching

► Footwear Recognition

Overfitting

The phenomenon that the learning result performs very good on training data but poorly on unseen new data, which is caused by that the learning approach has fit the training data too much, such that some malign particularities that prevent good generalization has also been captured by the learning result.

► Ensemble Learning
► Mainfold Learning

P

Paleoanthropology

Paleoanthropology is the study of ancient human beings, with the evidences such as bones, and footprints.

► Skull, Forensic Evidence of

Palm Dorsal Vein

► Hand Veins

Palm Segment

It refers to one of the three palm portions: lower palm, upper palm, and writer palm.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Palm Vein

Masaki Watanabe
Fujitsu Laboratories Ltd, Kawasaki, Japan

Synonyms

Palm vein authentication; Palm vein recognition

Definition

Palm vein authentication is one of a modality of biometric authentications, and is classified as a physiological biometric authentication. It uses palm vein patterns, a vascular image of a person's palm which can be seen as a kind of pattern, as personal information.

The palm vein patterns are normally captured under near-infrared illumination using the reflection method, in which ► near infrared rays are emitted from a person's palm and the reflected light is captured. The places on the palm where veins occur are captured as dark parts because veins absorb more near-infrared illumination while emitting only little. With this reflection method, a ► contactless type of pattern-capturing and user identification can be realized.

Because veins are inside the human body, they are secure and hard to be stolen or duplicated. Moreover, because palm vein patterns are varied and complex, they have sufficient information to identify one individual among many people and palm vein authentication is highly accurate.

A contactless type of identification is suitable for applications that require a high level of hygiene and for public-use applications. Several banks in Japan use palm vein authentication to identify customers since July 2004. In addition, the method has been used in a variety of applications including door security systems of offices and condominiums, login management systems for PCs, and to identify patients in hospitals.

Introduction

Palm vein authentication is one of the vascular pattern authentication technologies. It uses palm vein patterns, which are difficult to be seen by human on a person's palm, as personal information. Because palm vein patterns are information that are found within someone's body, it is hard for that information to be stolen.

© 2009 Springer Science+Business Media, LLC

This means that forgery is very difficult under usual conditions.

Palm vein patterns are unique to each individual; even identical twins have different palm vein patterns. Furthermore, palm vein patterns do not change within lifetime of a human except in certain cases of injury or disease. Although these facts and any other influences including physiological growth have not been medically proven, as with the fingerprints, irises, and other internal identification methods, experimental results based on extensive data and large-scale operational results prove practically that palm vein authentication has the advantages of consistency and accuracy as a method of personal identification.

A patent for hand vein authentication was filed in 1985 by Joseph Rice in the United States [1]. The first device for palm vein authentication was presented by Advanced Biometrics, Inc. in the United States in 1997, and in 2003, a remarkable contactless device was released by Fujitsu in Japan. In 2004, Japanese financial institutions adopted Fujitsu's technology for confirming the identification of its customers. This was the first major application in Japan in which a private enterprise adopted biometric authentication in a service for the general public. Palm vein authentication and finger vein authentication have received a great deal of attention in Japan compared to other biometric authentication methods such as fingerprint, iris, and face recognition methods.

Fujitsu's implementation of a contactless sensor and its concept was awarded the "Wall Street Journal's 2005 Technology Innovation Award for Security in Networks" [2].

Palm Vein

A person's palm has a wide and complex vein pattern, and it contains sufficient information to identify an individual among many people. If other parts of the hand are used for authentication, additional information, such as the relative position of that part of the hand relative to a vein sensor will be needed because the vein pattern of other parts of the hand does not contain sufficient information for identification.

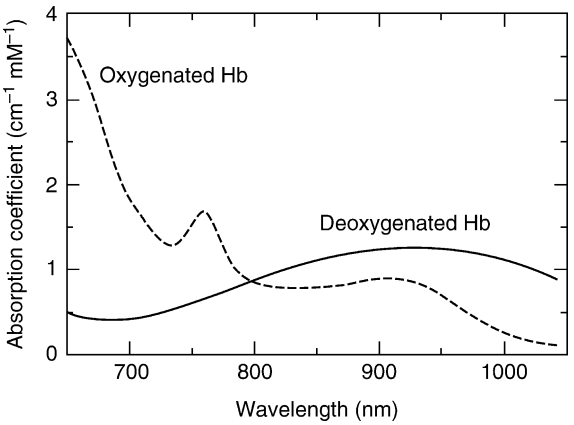
Compared with the back of the hand or the back of a finger, the palm is a good area to use because it does not normally have any hair on it to obscure the vein pattern.

Sensing

Vein patterns in the subcutaneous tissue of a person's palm are captured using near-infrared rays. This technology is called near-infrared spectroscopy (NIRS) and imaging. This has been investigated as a technology of in vivo measurements in the last ten or so years [3].

Hemoglobin is grouped into two types; (1) oxygenated hemoglobin that is present in arteries, which contains oxygen; and (2) deoxygenated hemoglobin present in veins that does not contain oxygen, and in particular it absorbs light with a wavelength of about 760 nm (Fig. 1) [4, 5]. When capturing an image of a palm using near-infrared rays which include the wavelength of light, the veins appear like a shadow on the palm, seen to be darker than the surrounding area (Fig. 2). In other experiments using near-infrared rays with a wavelength of 880 nm, a vein with a 1 mm diameter could be seen even if it is 3 mm below the surface of the skin [6].

Palm vein patterns are captured using a reflection method, which illuminates a person's palm from the front and also captures an image from the front of the palm. A palm image can be captured using light transmission method, which illuminates the palm from the back of the hand and captures an image from the front [6]. While in the transmission method the illumination device and the capture device are separated facing each other across a palm, in the reflection method, the illumination device and the capture device can be gathered together more compactly because the



Palm Vein. **Figure 1** Absorption spectra of hemoglobin. (Adapted from Wray et al. (1988) by K. Shimizu, Hokkaido University.)

Palm Vein. Figure 2 Infrared ray image of a palm.

direction of the illumination is the same as the direction of image capturing.

In Fujitsu's implementation [7, 8], a palm vein authentication sensor is made in the shape of a small box 35 mm deep × 35 mm wide × 27 mm high. Capturing is executed in a contactless manner. Users do not need to touch the sensor; they only have to hold their palms above it. To obtain a clear image of the palm vein pattern of a hand floating in the air, the capturing is controlled according to the movement and height of the hand above the sensor, and the illumination is controlled according to the light around the sensor.

In contactless manner, worries of the user regarding their sensitiveness about hygiene or any emotional concerns can be eliminated. It enables the identification method to be used in environments where a high standard of hygiene is required, such as in hospitals or food factories. In addition, sufficient consideration is given to individuals who are reluctant to come into direct contact with publicly used devices.

The intensity of the near-infrared ray emitted from the sensor is also safe. It is lessser than the intensity specified in the "Light and Near-Infrared Radiation" guidelines of the American Conference of Governmental Industrial Hygienists (ACGIH).

Matching

When first matching a palm to a database of images, a palm vein pattern is extracted from a near-infrared image. The pattern is picked up among the dark lines, which are obtained by morphologically tracing valleys of brightness in the palm area in an image.

The similarity of the captured vein pattern with one in a database is given a score to identify whether the vein patterns that have been registered are the same as that has been captured from a person being authenticated. This rating is based on the sum of the Euclidean distance between pixels that compose the two palm vein patterns.

For a verification (one-to-one matching), by way of example, the identity of the user is proved if the similarity score is greater than or equal to a predetermined threshold. Otherwise, the user is regarded as an imposter.

For an identification (one-to-many matching), similarity scores are calculated between the palm vein pattern captured from the person being identified and all or some palm vein patterns that have been registered in a database, and the identity of the palm vein pattern which has the maximum similarity score is regarded as being the one of the person being identified.

Performance

Using the data of 150,000 palms from 75,000 people [7], it is suggested that a typical palm vein authentication system can achieve false acceptance rate of less than 0.00008% and a false rejection rate of 0.01%, provided that the palm is held over the sensor two times during registration and one retry is allowed for comparison during authentication [7]. These results are certainly promising and confirm the individuality of finger vein features in a large user population.

In addition, ability of the sensor to perform personal authentication was verified using the following data: (1) data from individuals ranging from 5 to 85 years old, including people in various occupations, in accordance with the demographics released by the Statistics Center of the Statistics Bureau in Japan; (2) data from foreigners living in Japan in accordance with world demographics released by the United Nations; (3) data that trace daily changes in the palm vein pattern over several years; and (4) data taken in various situations in daily life, for example after drinking alcohol, taking a bath, going outside, or waking up.

Palm vein authentication technology was evaluated in Round 6 of Comparative Biometric Testing (CBT) by International Biometric Group (IBG) in 2006. CBT of IBG evaluates the accuracy and usability of biometric products using scenario-based testing, and strives to understand biometric performance in the real-world conditions. CBT Round 6 was the first major independent test to evaluate multiple vascular recognition technologies. Such assessments are typically based on a comparison of recognition samples and enrollment templates. In the case of palm vein authentication, approximately 40,000 [8] genuine comparisons and 50 million imposter comparisons were executed.

Results of the IBG study revealed that palm vein authentication performed exceptionally well in the failure to enroll (FTE) testing; only one person out of 1,290 did not finish the enrollment process given the test criteria, a failure rate was of only 0.08%. This extremely low rate indicates that palm vein authentication is highly applicable virtually for every individual, and does not impose any physiological restrictions when users interface with the device. This further indicates that palm vein authentication is usable, is easy for the users to learn, and is ideal for use in high-volume and large-scale applications.

Most importantly, palm vein authentication was effective when tested for authentication accuracy. The false acceptance rate (FAR) and false rejection rate (FRR) were extremely low, outperforming other products in the evaluation at standard and high security. The performance differences between same- and different-day transactions were also minimal. Therefore, after the users learned how to use the device, they were able to use it successfully on an ongoing basis. These data further confirm that palm vein authentication is highly accurate and has optimal usability, both of which are relevant to real-world conditions.

Implementation

In actual implementation, palm vein patterns can be stored on a smartcard. The matching between the palm vein pattern on the smartcard and the one captured for authentication can also be executed on the smartcard. Because the palm vein pattern is protected against external attacks by an antitampering function of the smartcard, users can handle their own palm vein patterns safely.

Application

Door Security Systems

Palm vein authentication sensors have been installed on many access control units over the world (Fig. 3). They can be used to control entry and exit into and out of rooms and buildings. For those applications, the combination of the following features of palm vein authentication means that it provides the optimum system: a hygienic contactless unit ideal for use in public places, user-friendly operation that requires people to simply hold their palms over a sensor, and a method that makes impersonation difficult.

In view of the Personal Information Protection Act that went into full effect in Japan in April 2005, the Department of Planning, Information and Management of the University of Tokyo Hospital began using palm vein authentication for a new security system to control room access. The security levels of the system were divided into three access levels: access to the administrative room, the development room, and the server room. A palm vein authentication access control unit has been installed at the entrance of each room. The system has been able to restrict an individual's entry in stages.

Palm Vein. Figure 3 Palm vein access control unit implemented by Fujitsu.

Palm Vein. **Figure 4** PC mouse containing palm vein authentication sensor.

Login Authentication

Palm vein authentication sensors can be integrated into PC mouse (Fig. 4). Using a mouse as a palm vein authentication sensor offers convenience and space-saving advantages.

Many companies and government agencies have an internal information system which handles sensitive personal data. Using a mouse with an integrated palm vein authentication sensor enables advanced, high-level security for system log-ins, beyond mere IDs and passwords, with the high accuracy and reliability of palm vein authentication.

Palm Vein. **Figure 5** ATM with palm vein authentication sensor.

Financial Services

In 2003, Japan saw a rapid increase in financial damage caused by fraudulent withdrawals from bank accounts by spoofing the identity with fake bankcards that were made using information from stolen or skimmed cards. It was a significant social problem. This had caused a sharp increase in the number of lawsuits taken out by victims against financial institutions for their failure to control information used for personal identification. The "Act for the Protection of Personal Information" came into effect in May 2005, and in response, financial institutions in Japan have been focusing on biometric authentication methods together with smartcards, as a way to reinforce the security of personal identification. Palm vein authentication is the form of biometric authentication that was most quickly introduced for customer confirmation at banking facilities. It has been used since July 2004, before the act came into effect.

When used for financial services, a user's palm vein pattern is registered at a bank counter and stored on a smartcard. This has the advantage of allowing users to carry their own palm vein pattern around with them,

and lets them manage the usage of their smartcard. To verify ATM transactions, palm vein pattern of a user's is captured by a palm vein authentication sensor on the ATM (Fig. 5). The captured palm vein pattern is transferred to the user's smartcard, and this with the one transferred from the sensor are compared on the smartcard. Finally, only the matching result and not the registered palm vein pattern of the user is output from the smartcard.

In addition to Japan, Brazil has also already decided to adopt palm vein authentication to identify users in ATM banking transactions. Banco Bradesco S.A., the largest private bank in Latin America, has been on testing palm vein authentication. After researching various biometric technologies, Bradesco chose palm vein authentication because of its outstanding features, such as its high level of verification accuracy and the fact that it is noninvasive and hygienic, making it easier to be accepted by customers of the bank.

Healthcare

Palm vein authentication is being deployed throughout the Carolinas HealthCare System (CHS) in the United States as part of a solution to effectively register patient information and ensure that the proper medical care is given to the right person, while protecting their medical record and privacy from identity theft and insurance fraud. To implement the system, CHS developed a unique hand guide as a sensor. The hand guide is adapted perfectly for a hospital environment since it incorporates a pediatric plate that adapts the guide so it can be used with young children, and can accommodate all the patients of the CHS.

The Sapporo Hospital of Keiyu Association in Japan also adopted palm vein authentication for their electronic medical records system for patient authentication. Patients who are to undergo an operation, register their palm vein patterns before the operation and the registered palm vein pattern and the palm vein pattern scanned from the patient on the day of the operation are compared. This confirms that the patient is the same as the one whose records have been input in the electronic medical recording system by the doctor in charge, and avoids the wrong patient being operated on, which might occur if two patients have the same name for example.

Some applications for healthcare could be realized because the contactless type of palm vein authentication is excellent in terms of hygiene.

Other Uses

The Chiba Institute of Technology in Japan deployed a student ID system that combines palm vein authentication and multifunctional smartcards to verify the identity of students, and lets them securely access their academic transcripts and other personal records through information kiosk terminals installed in various locations around the campus.

An examination service can use palm vein biometrics to authenticate the identity of examination candidates. Palm vein authentication is viewed as preferable to other modalities due to the reliability; it is not easy to steal palm vein images of others, making spoofing difficult.

Summary

Palm vein authentication uses vein patterns on the palm of a person as personal information. It is a highly secure technology because palm vein pattern is information contained within the body of someone. It is also highly accurate because palm vein patterns are complex and unique to each individual. Moreover, its contactless feature gives it a hygiene advantage over other authentication technologies. Many users of practical applications have highly evaluated this authentication method and experienced no psychological resistance to using it. This is good reason for developing new products for various solutions, starting with financial solutions followed by access control units and then login sensors.

Related Entries

- ▶ Vascular Image Data Format, Standardization
- ▶ Vascular Network Pattern
- ▶ Vein

References

1. Rice, J.: Apparatus for the Identification of Individuals. US Patent **4**,699,149 (1985)

2. Totty, M.: A better idea. Wall St. J. Oct. 24. (2005)
3. Kim, J.G., Xia, M., Liu, H.: Extinction coefficients of hemoglobin for near-infrared spectroscopy of tissue. IEEE Eng. Med. Biol. Mag. **24**, 118–121 (2005)
4. Wray, S., Cope, M., Delpy, D.T., Wyatt, J.S., and Reynolds, E.O.: Characterization of the near infrared absorption spectra of cytochrome aa3 and haemoglobin for the non-invasive monitoring of cerebral oxygenation. Biochim. Biophys. Acta. **933**(1), 184–192 (1988)
5. Cope, M.: The application of near infrared spectroscopy to non invasive monitoring of cerebral oxygenation in the newborn infant. Ph.D. thesis, University College London, Appendix B 316–323 (1991)
6. Editorial Board for Visualization: Techniques of Biological Information Visualization Techniques of Biological Information (in Japanese), Corona Publishing Co., Ltd., p. 86 (1997)
7. Watanabe, M., Endoh, T., Shiohara, M., Sasaki, S.: Palm vein authentication technology and its applications, In: Proceedings of Biometrics Symposium. 37–38 (2005)
8. International Biometric Group: Comparative Biometric Testing Round 6 – Public Report. http://www.biometricgroup.com/reports/public/reports/CBT6_report.htm. Accessed 19 Mar, 2009
9. Sasaki, S., Hiroaki Kawai, H., Wakabayashi, A.: Business expansion of palm vein pattern authentication technology, FUJITSU Sci. Tech. J. **41**(3), 341–347 (2005)

Palm Vein Authentication

► Palm Vein

Palm Vein Authentication Sensor

► Palm Vein Image Sensor

Palm Vein Image Sensor

Masaki Watanabe
Fujitsu Laboratories Ltd., Kawasaki, Japan

Synonyms

Palm vein authentication sensor; Palm vein scanner

Definition

The palm vein image sensor is used for palm vein authentication. The device captures an image of the vein pattern in the palm by emitting near-infrared rays that are absorbed by the deoxygenated hemoglobin in the veins and then reflected back to the device for image capturing.

The palm vein image sensor is commercially available as a device for more secured personal identification. Example applications include door security systems, PC login management systems, financial services security systems, and hospital patient confirmation systems.

Introduction

The palm vein image sensor is used for palm vein authentication, a vascular pattern recognition technology. The device uses the vein pattern of the palm as personal identification data. Therefore, the palm vein image sensor must scan the position of the veins with the highest degree of accuracy.

The technology for noninvasive scanning of blood vessels is primarily categorized as in vivo measurement. In this field, near-infrared spectroscopy (NIRS) and imaging has been investigated for the last ten years [1]. The palm vein image sensor was developed based on this NIRS technology.

Infrared scanning of the pattern of subcutaneous blood vessels for the identification of individuals was first disclosed in Rice's patent in 1985 [2]. In 2001, Peterson et al. patented a device for palm vein authentication [3], in which a plurality of light-emitting elements is arranged in an array with light-detecting elements in a flexible mat. The mat was intended to be put into a device to be grasped by the hand of the person to be identified following to Stiver's patent [4]. The device comprises of an elongated transparent cylindrical exterior shell and scans the hand grasping the device. In 2003, Fujitsu realized a box-type device for practical use and it was launched in Japan the following year. In 2008, Snowflake Technologies [5] released a prototype in which a palm vein sensor is embedded to the upper part of equipment for a door security system. The palm vein image patterns can also be simultaneously acquired with palmprint images and employed to achieve improvement in performance of palmprint authentication as detailed in [12].

Hemoglobin in vessels is grouped into two types: oxygenated hemoglobin that is in arteries and contains oxygen; and deoxygenated hemoglobin that is in veins and does not contain oxygen, and in particular, absorbs light with a wavelength of about 760 nm [6, 7]. When the palm of the hand is illuminated by near-infrared light, the rays will be scattered by structures under the surface of the skin. Much of this illumination will be reflected back towards the illumination source, however little will be reflected from the veins since these absorb near-infrared.

Palm vein patterns are preferably acquired using the reflection method [3] and also in commercially available products from Fujitsu [9] and Snowflake [5], whereby the palm is illuminated from the front and the image is also captured from the front. If the transmission method is used, whereby the palm is illuminated from the back of the hand and the image is captured from the front, a very strong light would be needed.

To realize this imaging method, the palm vein image sensor must have both of an illumination function and imaging function by near-infrared rays. The illumination function must emit light in a wavelength of about 760 nm and the imaging function must have sufficiently high resolution to distinguish the vein pattern. In the reflection method, because the direction of illumination is the same as that of image capturing, the illumination device and the imaging device are compactly integrated.

A palm vein image sensor should ideally scan as broad a palm area as possible for keeping high accuracy of palm vein authentication because the human palm has an extensive and complex vein pattern that contains sufficient information to identify an individual from among many people. But some kinds of sensors which scan a partial area of a palm would be designed for reasons of usability such as the Stiver's patent.

Implementation

The research and development efforts for the compact palm vein image sensors have been confined to few commercial vendors. Therefore only very limited technical details are available for the palm vein image sensors. In Fujitsu's implementation [8, 9], the palm vein image sensor is in the shape of a small box 35 mm deep by 35 mm wide by 27 mm high. Image capturing is executed in a contactless manner. Users do not touch the sensor; they only have to hold their palms above it. The user places his or her hand below an optical reader, which scans the palm.

To obtain a clear image of the palm vein pattern, imaging is controlled according to the movement and position of the hand above the sensor or below the sensor, and illumination is controlled recognizing the light around the sensor. Video-rate scanning is typically employed for the convenience in the palm vein authentication so that users do not have to stop the hand for authentication.

Any imaging devices such as CMOS sensors or CCD sensors will be used for the capturing but it must have sensitivities of near-infrared rays. It should be also assembled not to capture except for the reflected near-infrared rays from inner of the hand using such as a polarizing filter, an optical filter cutting off visible lights, and so on.

The contactless method, which is adopted by both of sensors, eliminates the concerns of users who are sensitive about hygiene or who are reluctant to come into direct contact with publicly used devices. It also enables the identification method to be used in environments where a high standard of hygiene is required, such as in medical facilities or food factories.

Regarding security, a data encryption function for the palm image should be also provided. This ensures that image output is protected from any unauthorized access or tampering.

Other possible substantiation of palm vein image sensor has been also announced. NEC Corporation developed the world's first contactless multi-modal finger recognition technology [10]. The new device quickly scans two forms of biometric information, fingerprints and vein patterns of finger. They proved that the device might also be adapted to recognize skin and vein patterns from any region of a human body. As an example, the possibility of a scanning system customized to analyze and authenticate both fingerprint and palm characteristics was also shown.

Evaluation

Comparative Biometric Testing (CBT) by the International Biometric Group (IBG) evaluates the accuracy and usability of biometric products using

scenario-based testing, and strives to understand biometric performance in real-world conditions. Round 6 of the testing in 2006 evaluated palm vein authentication technology using Fujitsu's sensor.

The results of the IBG study revealed that palm vein authentication using this sensor performed exceptionally well in the failure to enroll (FTE) testing [11]. Authentication accuracy was also good; the false acceptance rate (FAR) and false rejection rate (FRR) were extremely low. Performance differences between same-day and different-day transactions were also minimal. Thus, after the users learned how the sensor worked, they were able to use it successfully on an ongoing basis. This data further confirms that palm vein authentication using this sensor is highly accurate and has optimal usability, both of which are directly relevant to real-world conditions.

Palm Vein Image Sensor. **Figure 1** Palm vein access control unit implemented by Fujitsu.

Application

Palm vein image sensors are embedded in many different types of equipment for various applications, such as for door security systems, PC login, financial services, and so on.

The access control unit for door security systems is a typical example in which the palm vein image sensor is installed. Units throughout the world are equipped with this sensor using various methods. Almost all the units have an input device for the user's ID, such as a ten-key, smartcard reader, or both (Fig. 1).

Moreover, for the purpose of PC login, the palm vein image sensor can be installed in the PC mouse (Fig. 2), and in the case of financial services, it can be installed in almost any type of ATM (Fig. 3).

Some of the equipment installed with a palm vein image sensor includes a hand guide on which the wrist is placed, so that first-time users of palm vein authentication can easily understand how to use this sensor. The adaptation of the Carolinas HealthCare System (CHS) in the United States, which

Palm Vein Image Sensor. **Figure 2** PC mouse equipped with palm vein authentication sensor.

Palm Vein Scanner. Figure 3 ATM with palm vein authentication sensor.

is designed to protect patients from identity theft and insurance fraud, is a good example of utilizing the hand guide. To implement the system, CHS developed a unique hand guide, adapted perfectly for a hospital environment since it incorporates a pediatric plate for young children, and thus can accommodate all CHS patients.

Summary

The palm vein image sensor is used for palm vein authentication employing the vein pattern of the palm of a person as personal identification information. An image of the palm vein pattern is captured using near-infrared rays in the reflection method. The illumination device emits light with a wavelength of around 760 nm, which is more strongly absorbed by the deoxygenated hemoglobin in the veins compared to the surrounding subcutaneous tissues of the hand. The imaging device captures the whole or partial palm area and has a sufficiently high resolution to distinguish the vein pattern. In case of sensors using the reflection method for capturing, those two devices are compactly installed. The palm image sensor is used for various applications and is embedded in many different kinds of equipment, such as the access control unit for door security systems, PC mouse for login management, ATMs for financial services, confirmation units for hospital patients, and so on.

References

1. Kim, J.G., Xia, M., Liu, H.: Extinction coefficients of hemoglobin for near-infrared spectroscopy of tissue, IEEE Eng. Med. Biol. Mag. March/April, 118–121 (2005)
2. Rice, J.: Apparatus for the Identification of Individuals, US Patent, 4, 699, 149 (1985)
3. Peterson, D.C., Jackson, D.W., Stiver, J.A: Method and apparatus for subcutaneous identification, US Patent, 6,330,346 (2001)
4. Stiver, J.A. Peterson, D.C.: Identification system, US Patent, 5,793,881 (1998)
5. Snowflake Technologies, http://snowflaketechnologies.com/ Accessed 19 Mar, 2009
6. Wray, S., Cope, M., Delpy, D.T., Wyatt, J.S., Reynolds, E.O.: Characterization of the near infrared absorption spectra of cytochrome aa3 and haemoglobin for the non-invasive monitoring of cerebral oxygenation, Biochim. Biophys. Acta, **933**(1), 184–192 (1988)
7. Cope, M.: The application of near infrared spectroscopy to non invasive monitoring of cerebral oxygenation in the newborn infant, Ph.D. thesis, University College, London Appendix D. 316–323 (1991)
8. Watanabe, M., Endoh, T., Shiohara, M., Sasaki, S.: Palm vein authentication technology and its applications, In: Proceedings of Biometrics Symposium, pp. 37–38 (2005)
9. Sasaki, S., Hiroaki Kawai, H., Wakabayashi, A.: Business expansion of palm vein pattern authentication technology, FUJITSU Sci. Tech. J. **41**(3), 341–347 (2005)
10. NEC Corporation: NEC develops the world's first contactless multi-modal finger recognition technology (2008), http://www.nec.co.jp/press/en/0805/1402.html. Accessed 19 Mar, 2009
11. International Biometric Group: Comparative Biometric Testing Round 6 – Public Report (2006), http://www.biometricgroup.com/reports/public/reports/CBT6_ report.htm. Accessed 19 Mar, 2009
12. Wang, J.-G. Yau, W.-Y. Suwandy A. Sung, E. Person recognition by palmprint and palm vein images based on 'Laplacianpalm' representation. Pattern Recogni. **41**(5), 1531–1544 (2008)

Palm Vein Recognition

▶ Palm Vein

Palm Vein Scanner

▶ Palm Vein Image Sensor

Palmprint

Palmprint is an impression or image left on a surface by the friction skin of the palm.

► Anatomy of Friction Ridge Skin

Palmprint, 3D

David Zhang, Vivek Kanhangad
Biometrics Research Centre, Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Synonyms

Palmprint authentication, 3D; Palmprint recognition, 3D

Definition

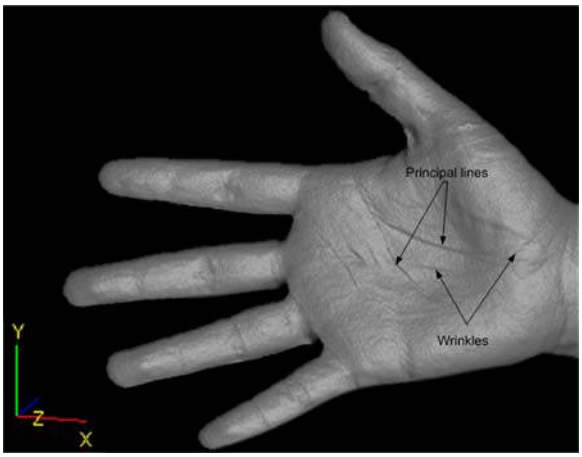
Biometric systems that rely on unique features of the palm for personal recognition are referred to as palmprint based biometrics. A 3D Palmprint biometric system employs 3D imaging device to acquire surfaces of human palm and uses this data in performing user identification. Extracted features from a 3D palmprint data include depth and curvature of palmlines and wrinkles on the palm surface.

Introduction

Human palmprints are rich in features that are unique and stable. Major palmprint features include principal lines, wrinkles, ridges, singular points, and minutiae points (see Fig. 1). In addition, human palmprints are also abundant with texture features. Apart from being feature-rich, palmprints have advantages over other hand-based biometric technologies [1, 2]:

1. Compared to the fingerprint, the palm provides a larger surface area so that more features can be extracted
2. An individual is less likely to damage a palm than a fingerprint, and the line features of a palm are stable throughout one's lifetime
3. Small amounts of dirt or grease appearing on an individual's finger may pose challenges in accurately extracting features for a fingerprint system. This problem does not arise in the extraction of palmprint features since a comparative low resolution palmprint images are used to extract these features

Most of the current works in the area of automated palmprint recognition are based on acquiring intensity image of the user's palm and extracting line or texture features from it. Although these systems have been able to achieve promising performance with low error rates [3], there are a few inherent limitations associated with such systems. Firstly, they are sensitive to spurious patterns such as dirt, lines or text on the palm. Performance of the palmprint recognition systems based on intensity images can be severely affected by the spurious patterns on the palm. An impostor may also blemish his palm with a purpose to circumvent the system. In addition, like most other 2D image based personal recognition systems, palmprint systems are also vulnerable to sensor level spoof attacks. An impostor may easily fabricate a spoof palm resembling a genuine user's palm and use to it to circumvent the system. On the other hand, a palmprint biometric system based on 3D images of users' palm offers higher degree of robustness against such attacks. These systems are extremely difficult to circumvent as they require sophisticated methods to fabricate spoof 3D palmprint models. Another factor that often has an impact over the performance of



Palmprint, 3D. Figure 1 Palmprint features.

conventional palmprint recognition systems is the changes in the illumination. This, in fact is a major problem for other biometric systems such as the one employing users' face images. The face images are often required to be imaged in outdoor environments and, as a result, show large intra class variations. However, palmprint systems often acquire images in a controlled environment and therefore are less affected by inconsistent illumination.

One possible solution to overcome these limitations is to make use of depth features of the palm surface using a 3D imaging device. Such observations will provide information on depth and curvature of palmlines and wrinkles on the palm surface. Comparable performance and robustness make 3D palmprint biometric system a good candidate for high security identification tasks. However there has not been much research focused on exploring the utilities of 3D palmprint feature. Therefore references on this topic are limited.

3D Palmprint Recognition System

A 3D palmprint based personal recognition system includes the following modules:

1. 3D palmprint image acquisition device
2. Extraction of Region of Interest (ROI) to obtain the central part of the palmprint
3. Feature extraction
4. Feature matching, where extracted features are matched with their respective feature templates stored during enrolment phase, generating a similarity score. In identification applications, the query template is matched to all templates enrolled in the database. Therefore, a one-to-many comparison is performed in this case
5. Decision module, where the similarity score produced is compared to the threshold of the system to either accept or reject the identity claim. In identification applications, identity of the user is determined to be the one with highest similarity score

3D Palm Image Acquisition

Infrared sensors are employed to detect the presence of the hand on the acquisition device. When a hand is detected, the device projects multiple light patterns onto the palm surface and acquires depth information using active triangulation. In order to distinguish between stripes, they are coded with different brightness. The system uses a computer controlled liquid crystal display (LCD) projector that can generate arbitrary stripe patterns. A CCD camera is used to acquire the images formed on the palm side. The sequence of images acquired by the CCD camera is then processed to obtain the 3D palm data. Figure 2a shows the process of acquisition of 3D palmprint data using a 3D image acquisition device based on structured light principle. The US patent [4] describes the process of acquisition of 3D finger and palmprint information using multi-camera and light projection system. This device projects multiple structured lights from different directions and the images formed on the object (palm or finger) are captured by cameras at different angles. It is claimed that the system can simultaneously obtain 3D fingerprint and palmprint information.

Feature Extraction

Acquired 3D images are processed to extract the region of interest (ROI). The inter finger points are used as reference points to extract a sub image of fixed size located at the center part of the palm. These 3D sub images are further processed to extract surface curvature features. To represent the curvature of every point on the 3D palmprint image by a scalar value, the curvedness (C) introduced in [5] is utilized. The positive value C is a measure of how sharply or gently curved, a point is [6]. It is defined in terms of principal curvatures k_1, and k_2, as:

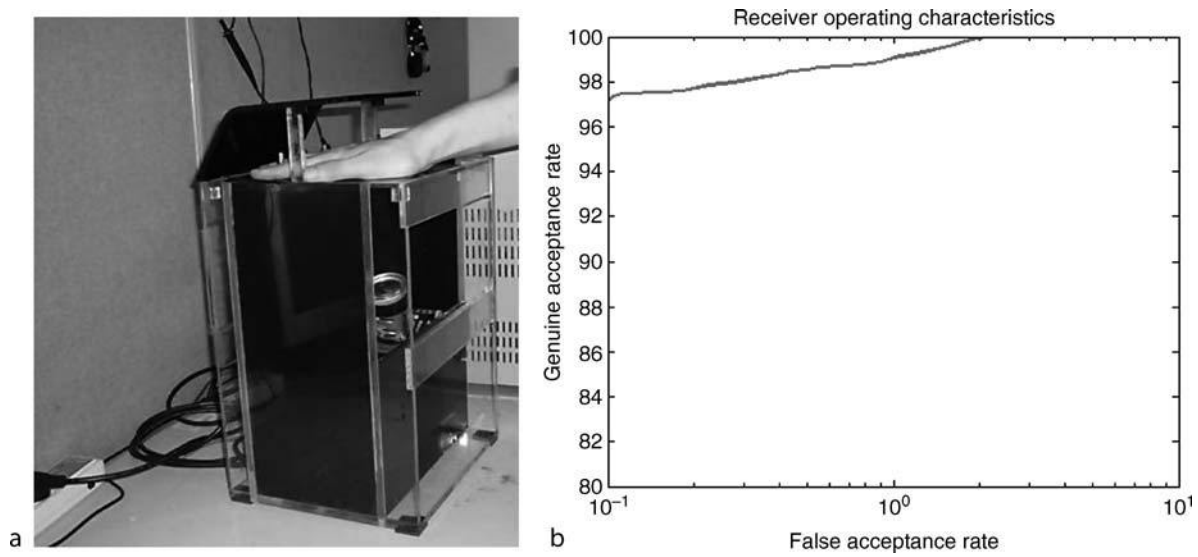
$$C = \sqrt{(k_1^2 + k_2^2)/2} \tag{1}$$

The principal curvatures k_1 and k_2 can be determined as:

$$k_1, k_2 = H \pm \sqrt{H^2 - K} \tag{2}$$

where K and H are Gaussian and mean curvatures respectively. For a surface patch represented by $X(u, v) = (u, v, f(u, v))$, the values of K and H are computed as follows:

$$K(X) = \frac{f_{uu} f_{vv} - f_{uv}^2}{(1 + f_u^2 + f_v^2)^2}$$



Palmprint, 3D. Figure 2 **(a)** Example 3D palmprint image acquisition system. **(b)** Receiver Operating Characteristics (ROC) curve for a 3D palmprint authentication system.

and

$$H(X) = \frac{(1+f_u^2)f_{vv} + (1+f_v^2)f_{uu} - 2f_u f_v f_{uv}}{(1+f_u^2+f_v^2)^{3/2}} \quad (3)$$

where, f_u, f_v and f_{uu}, f_{vv}, f_{uv} are first and second order partial derivatives of $f(u,v)$.

The scalar value of curvature (C) is obtained for every point on the 3D palmprint image and this can be stored in a 2D matrix or an image. Set of such scalar values is referred to as surface curvature map. Figure 3 shows sample 3D palmprint images and corresponding curvature maps. It can easily be observed that the surface curvature maps closely resemble the palmlines, especially the strong principal lines.

Additional features such as surface types that are characterized using the sign of mean and Gaussian curvature [7], shape index [8], computed using minimum and maximum curvature values; may also be explored as feature representations for the 3D palmprint images. Shape index, a local feature computed at each point, is defined as:

$$S = \frac{1}{2} - \frac{1}{\pi}\arctan\frac{k_1+k_2}{k_1-k_2} \quad (4)$$

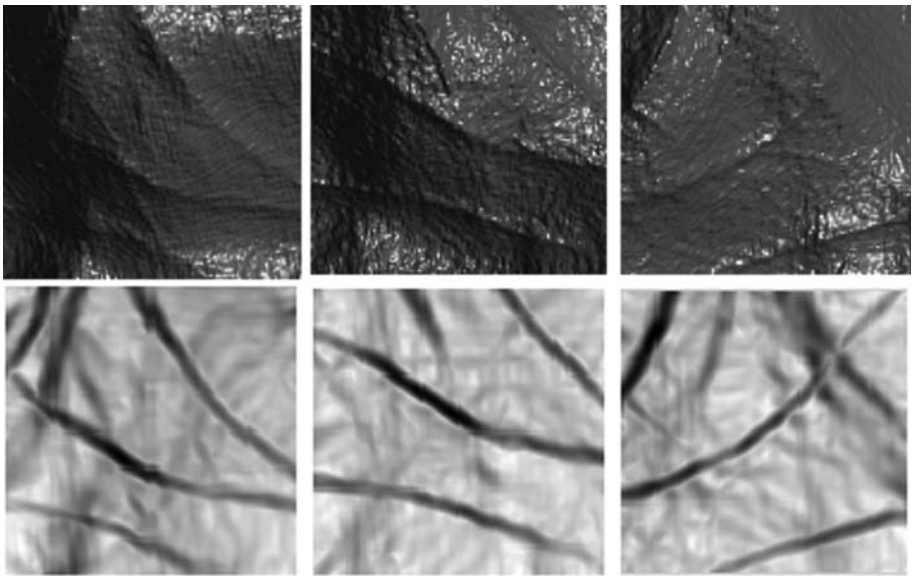
The value of shape index S lies in the interval [0,1] and can be used to classify each point on the surface to different surface types ranging from spherical cup to spherical cap.

Another approach to perform user authentication is to extract 3D palmlines from the acquired 3D palmprint images and match these lines (set of points) using point set alignment algorithms such as iterative closest point (ICP) [9]. ICP can be employed to iteratively estimate the transformation between the two 3D palm lines. Alignment error (e.g., mean squared distance) generated by the ICP can be used as the matching score.

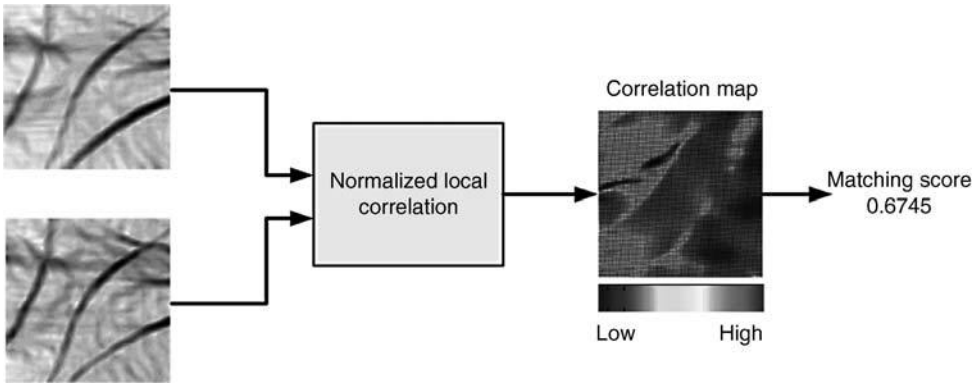
Feature Matching

Feature matching establishes the similarity between two 3D palmprint images. An image matching technique, normalized local correlation can be employed to compare two curvature maps. Result of this matching is a correlation value for every point in the input curvature maps. Average of these correlation values is considered to be the matching score. The expression for normalized local correlation is given by:

$$C = \frac{\sum_{i=-N}^{N}\sum_{j=-N}^{N}(P_{ij}-\bar{P})(Q_{ij}-\bar{Q})}{\sqrt{\left[\sum_{i=-N}^{N}\sum_{j=-N}^{N}(P_{ij}-\bar{P})^2\right]\left[\sum_{i=-N}^{N}\sum_{j=-N}^{N}(Q_{ij}-\bar{Q})^2\right]}} \quad (5)$$



Palmprint, 3D. **Figure 3** Sample 3D palmprint images (top row) and their corresponding curvature feature maps (bottom row).

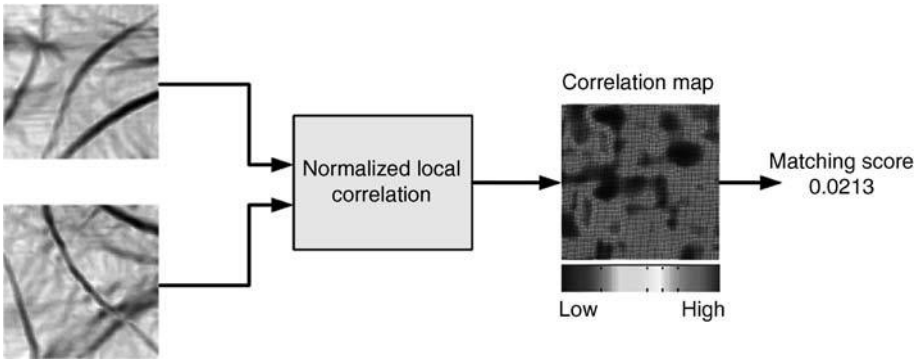


Palmprint, 3D. **Figure 4** Matching of two curvature maps from the same subject.

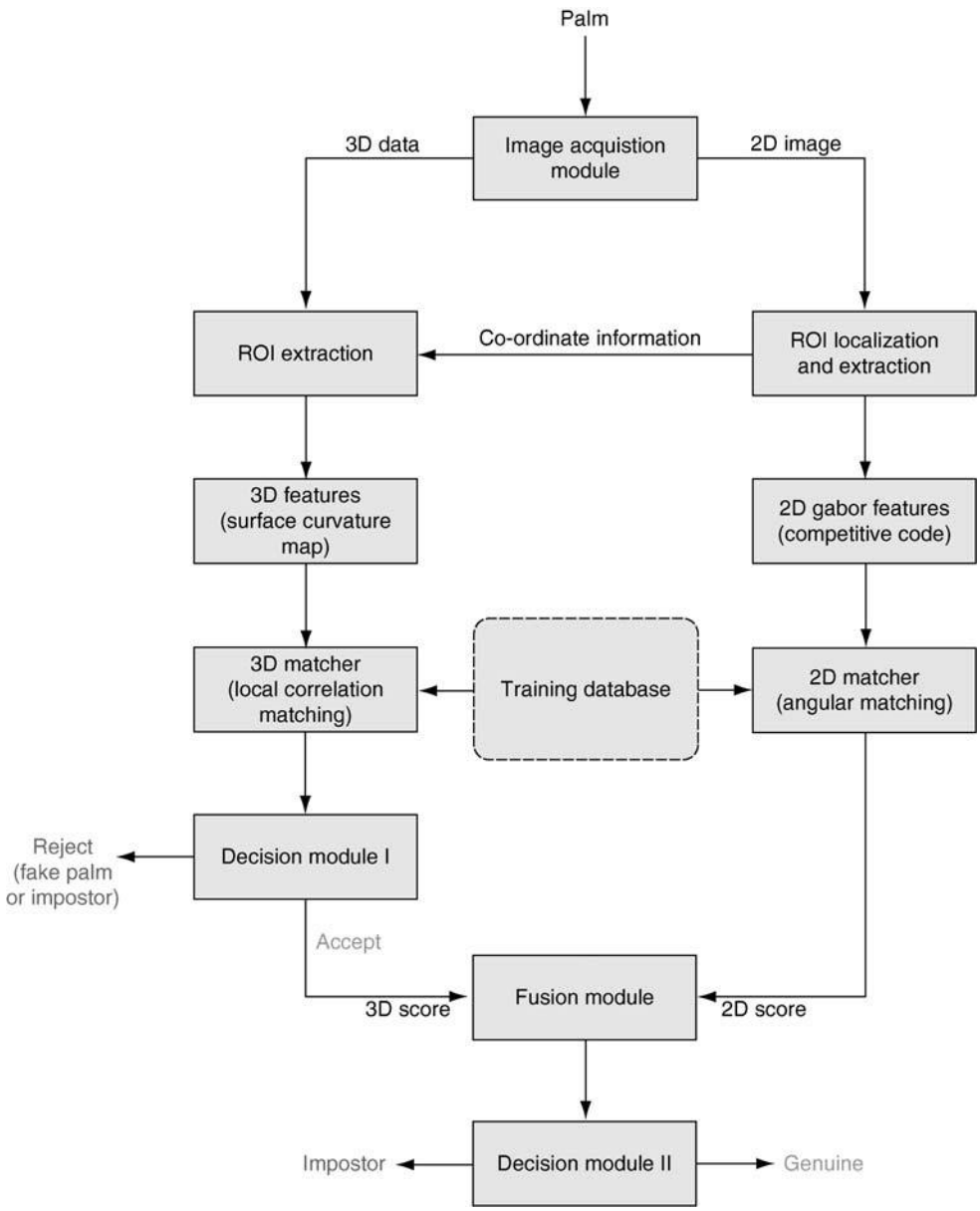
where P_{ij} and Q_{ij} are curvature values in the neighborhood of the points being matched in the two curvature feature maps, and $\bar{P}$ and $\bar{Q}$ are the mean curvature values in those neighborhoods. Figures 4 and 5 illustrate the process of matching two curvature maps of the same and different user respectively. Red (dark) colored pixels in the correlation map represent high values of correlation while blue (light) represents low correlation. Final matching score is the average of pixel values in the correlation map. It can be observed from Figure 4 that genuine matching results in a correlation map with large regions of red colored pixels, indicating high correlation between the two curvature maps being matched.

Decision Module

At the decision stage, match score from the feature matching process is used to make a decision as to whether the claimant is a genuine user or an impostor, (in the verification scenario) or to decide the identity



Palmprint, 3D. Figure 5 Matching of two curvature maps from different subjects.



Palmprint, 3D. Figure 6 A framework for combing 2D and 3D palmprint matching scores.

of the user in identification applications. Decision threshold is selected based on a threshold criterion, which is often the equal error rate (EER) of the system. Figure 6 presents a framework for combining 2D and 3D palmprint matchers. It employs a two level decision strategy. Decision module I accepts or rejects a claim based only on the 3D matching score. This allows the system to reject fake palms based on their 3D matching scores. At the second level, Decision module II operates on the combined (2D and 3D) matching score to achieve performance improvement.

Performance Evaluation

The 3D palmprint verification system developed in [10] presents promising performance on a database of 108 users. Six images were collected from each user, resulting in 648 3D palmprint images in the database. Matching each palmprint image with all other images in the database, 1,629 genuine and 208,008 impostor match scores were obtained. Figure 2b shows the receiver operating characteristic (ROC) curve for an authentication system based on 3D palmprint features. The reported system achieves an EER of 0.99% on the aforementioned database.

Summary

Palmprint based recognition systems are extensively researched for applications such as physical access control, attendance tracking, and other personal verification tasks. They have certain advantages such as high user acceptance and ease of use, over other biometrics. While 2D image based palmprint based recognition systems have achieved high performance, they can be sensitive to factors such as spurious patterns on the palm and variations in the illumination. 3D palmprint based recognition systems, on the other hand, offers high degree of robustness along with comparable performance, making them a good choice for high security applications. In addition, since 2D and 3D palmprint images can be simultaneously acquired using a single image acquisition device, these features can easily be combined to form a multi-biometric system that can potentially achieve significantly higher performance than either of the two palmprint features. The reference [10] provides details of a device that can simultaneously acquire 2D and 3D palmprint images. It also provides experimental results on score level fusion of 2D and 3D palmprint matchers. While 3D palmprint based biometric system has several advantages over other biometrics, size of these systems, due to large capture area, can be prohibitive for its use in devices like PDA and laptop.

Related Entries

- ▶ Hand Geometry
- ▶ Palmprint Feature
- ▶ Palmprint Matching

References

1. Zhang, D.: Palmprint Authentication, Kluwer, Dordrecht (2004)
2. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. Special Issue on Image- and Video-based Biometrics. IEEE Trans. Cir. Syst. Video Technol. **14**(1), 4–20 (2004)
3. Zhang, D., Kong, W.K., You, J., Wong, M.: On-Line palmprint identification. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1041–1050 (2003)
4. Chen, F. (2006) 3D Fingerprint and palm print data model and capture devices using multi structured lights and cameras. U.S. Patent No. 20060120576 (2006)
5. Koenderink, J.J., van Doorn, A.J.: Surface shape and curvature scales. Image Vis. Comput. **10**(8), 557–564 (1992)
6. Cantzler, H., Fisher, R.: Comparison of HK and SC curvature description methods. Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling, QC, Canada, pp. 285–291 (2001)
7. Besl, P.J., Jain, R.C.: Three-dimensional object recognition. ACM Comput. Surv. **17**(1), 75–145 (1985)
8. Woodard, D.L., Flynn, P. J.: Finger surface as a biometric identifier. Comput. Vis. Image Underst. **100**(3), 357–384 (2005)
9. Besl, P., McKay, N.: A method for registration of 3D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
10. Kanhangad, V., Zhang, D., Luo, N.: A multimodal biometric authentication system based on 2D and 3D palmprint features. Proceedings of the SPIE, 6944: 69440C–69440C-9 (2008)

Palmprint Anatomy

▶ Anatomy of Friction Ridge Skin

Palmprint Authentication, 3D

▶ Palmprint, 3D

Palmprint Characteristics

▶ Palmprint Features

Palmprint Database

▶ Hand Databases and Evaluation

Palmprint Device

▶ Fingerprint, Palmprint, Handprint and Soleprint Sensor

Palmprint Features

David Zhang, Laura L. Liu
Biometrics Research Centre, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Synonyms

Palmprint Representation; Palmprint Characteristics; Handprint

Definition

Palmprint features refer to the representation of palmprints which characterizes a palmprint in a stable and unique way such that it has good discriminating ability for personal identification. A palmprint identification system identifies an individual using palmprint features which may or may not be observable to the naked eye. The selection of palmprint features is a fundamental problem in reliable palmprint identification.

Introduction

Palmprint is the skin patterns of the inner surface of the human hand from the wrist to the root of the fingers. As a comparatively new biometric, palmprints are rich in physical characteristics of skin patterns such as lines, points and textures, which provide stable and distinctive information sufficient for separating an individual from a large population. Compared with other biometric traits, the advantages of palmprint are user friendliness, environment flexibility, and discriminating ability.

A typical palmprint-based identification system involves ▶ pre-processing, feature extraction, feature matching and decision-making. Normally, automatic palmprint identification systems can be classified into two categories: offline and online. An offline system usually processes previously captured palmprint images which are often obtained from inked palmprints or generated from ▶ forensic analysis, while an online system captures palmprint images using a palmprint scanner that is directly connected to a computer for real-time processing. For both types of palmprint identification systems, the selection of unique features to identify a person is a fundamental issue to be solved.

Many unique features of palmprint images can be used for personal authentication, which include ▶ principal lines, wrinkles, ridges, minutiae points, singular points, and texture. Various features can be extracted at different image resolutions. Features such as minutiae, ridges and singular points can be obtained in high-resolution palmprint images of at least 400 dpi (dots per inch) [1–3] and are difficult or even cannot be observed in low-resolution images (<100 dpi). Nevertheless, features like principal lines and wrinkles can be extracted from low-resolution palmprint images and play an important role in palmprint identification [4–6]. In the real-time palmprint identification system developed by the Biometric Research Centre at

The Hong Kong Polytechnic University [7], a very low resolution (75 dpi) palmprint image can be captured by a CCD (charge-coupled device) camera-based palmprint acquisition device (Fig. 1). In a typical palmprint image captured by this system, the main patterns can be generalized to principal lines, wrinkles, and creases (also called ridges) (Fig. 2).

Regardless of image resolution and applications, the selected feature should have good discriminating ability to exhibit large variations between individuals and small variations between samples from the same palm. Thus, image feature extraction/representation play an essential role in palmprint-based biometrics. In general, the palmprint features can be classified into three categories [8]: texture-, line-, and appearance-based features.

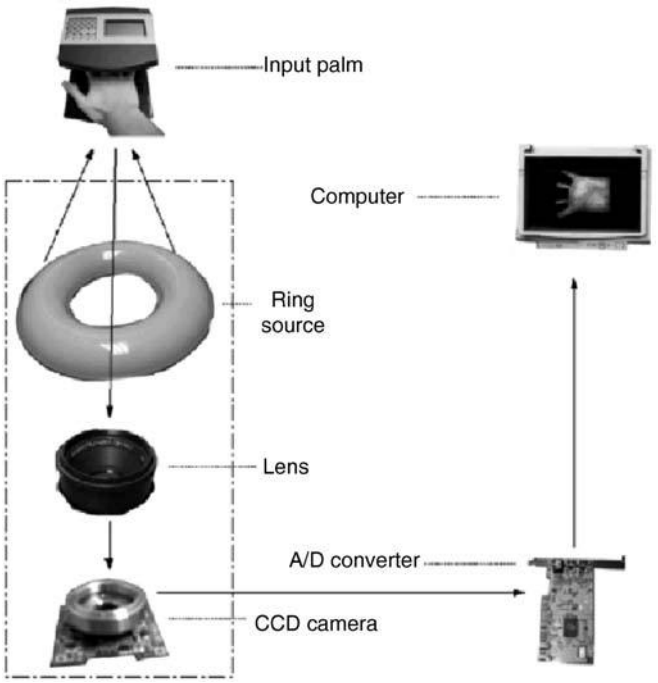
Texture Features

Texture representation of coarse level palmprint classification provides an effective approach to palmprint recognition. In the spatial domain, considering a palmprint as a texture image, a statistical approach, e.g., Laws' convolution masks, can be used to compute the texture energy of palmprints [2]. In the frequency domain, a polar coordination system (r, θ) can be established to represent the Fourier transform of palmprint images where less compact information indicates stronger line features in the spatial domain [9]. In addition, if a palmprint image has a strong line, there is more information along the line's perpendicular direction in the frequency domain. In this polar coordination system, the energy change tendency along r corresponds to the intensity of a palmprint's creases and that along θ to their directions. Another frequency domain approach to the palmprint feature extraction relies on the discrete cosine transform (DCT) of the image pixels [10].

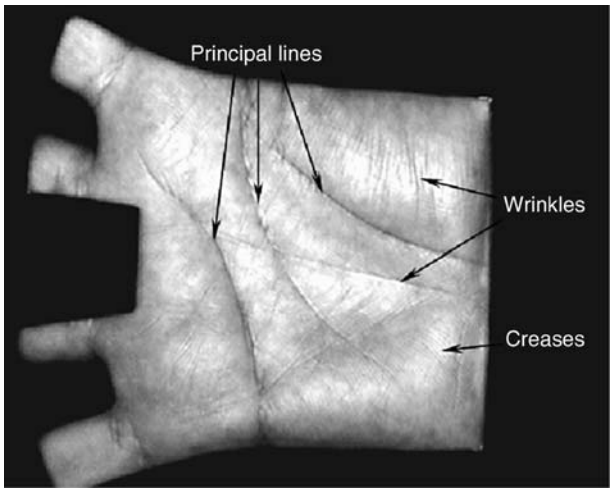
Two-dimensional Gabor filter, an effective tool for texture analysis, is widely used to extract texture features in palmprint-based biometrics [7, 11, 12]. A circular 2-D Gabor filter in the spatial domain has the following general form:

$$G(x, y, \theta, u, \sigma, \beta) = \frac{1}{2\pi\sigma\beta}\exp\left\{-\pi\left(\frac{x'^2}{\sigma^2}+\frac{y'^2}{\beta^2}\right)\right\} \exp(2iux'), \tag{1}$$

where $i = \sqrt{-1}, x' = (x - x_0)\cos\theta + (y - y_0)\sin\theta$, $y' = -(x - x_0)\sin\theta + (y - y_0)\cos\theta$, (x_0, y_0) is the



Palmprint Features. **Figure 1** The Design Principle of the Palmprint Acquisition Device.



Palmprint Features. **Figure 2** The Main Patterns in a Low-resolution Palmprint Image.

center of the function, u is the radial frequency in radians per unit length, and θ is the orientation of the Gabor function in radians. σ and β are the standard deviations of the Gassian envelop along x and y axes, respectively, and $\sigma = \beta$ for the circular 2-D Gabor filter. In order to provide more robustness to illumination, the Gabor filter is tuned to 0 DC (average value) with the application of the following formula:

$$\tilde{G}[x, y, \theta, u, \sigma, \beta] = G[x, y, \theta, u, \sigma, \beta] - \frac{\sum_{i=-n}^{n} \sum_{j=-n}^{n} G[i, j, \theta, u, \sigma, \beta]}{(2n+1)^2}, \quad (2)$$

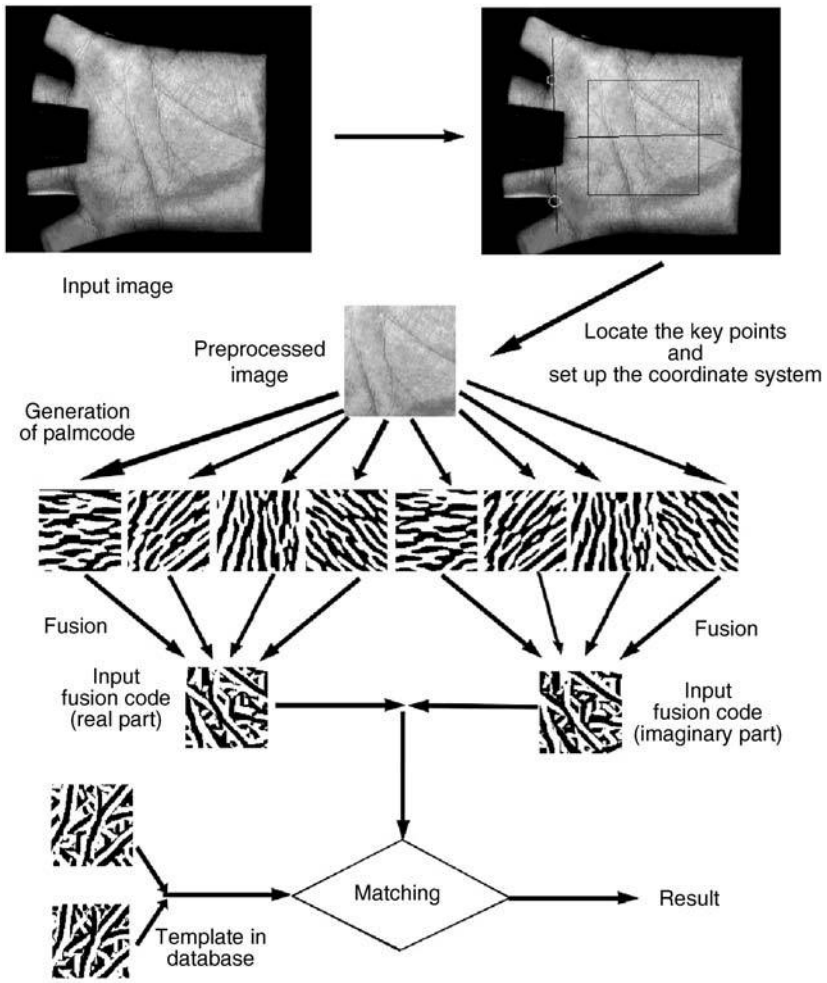
where $(2n+1)^2$ is the size of the filter and $G[i, j, \theta, u, \sigma, \beta]$ denotes the corresponding discrete Gabor filter. The adjusted Gabor filter is first used to convolute with the pre-processed central part palmprint sub-image and then each sample point in the filtered image is coded to two bits according to the signs of the real and imaginary parts of the convolution results. Using this coding method, only the phase information in palmprint images is stored in the feature vector, which is called the PalmCode.

The PalmCode has been improved in two aspects. To reduce the correlation between PalmCodes, a fusion rule has been developed to produce a single feature, called the Fusion Code. According to this fusion rule, multiple elliptical Gabor filters ($\sigma \neq \beta$) with different orientations are utilized to extract the magnitude and the phase information on a palmprint image and then employ the magnitude for fusion and the phase for the final feature (Fig. 3). To utilize the orientation information of palmprints which is absent in a PalmCode, a competitive rule is further designed to extract the orientation characteristics of palm lines by using a set of neurophysiology-based Gabor filters. This competitive rule is a winner-take-all rule where the orientation of the sample point is determined by the Gabor filter which gives the minimum filter response. The corresponding feature codes, referred to as the Competitive Code (Fig. 4), are more robust to different capturing environments and devices.

The ordinal feature representation is another powerful method to capture the texture features from low-resolution palmprint images [13]. In low-resolution palmprint images, the main palmprint patterns are negative line segments and can be characterized using ▶ ordinal measures. This measure qualitatively compares two elongated, line-like image regions which are orthogonal in orientation and generates one bit feature code. The 2-D Gaussian filter which is used to obtain the weighted average intensity of a line-like region is formulated as follows:

$$f(x, y, \theta) = \exp\left[-\left(\frac{x\cos\theta + y\sin\theta}{\sigma_x}\right)^2 - \left(\frac{-x\sin\theta + y\cos\theta}{\sigma_y}\right)^2\right], \quad (3)$$

where θ denotes the orientation of the 2-D Gaussian filter, σ_x and σ_y denote the filter's horizontal and vertical scales, respectively and σ_x/σ_y should be higher



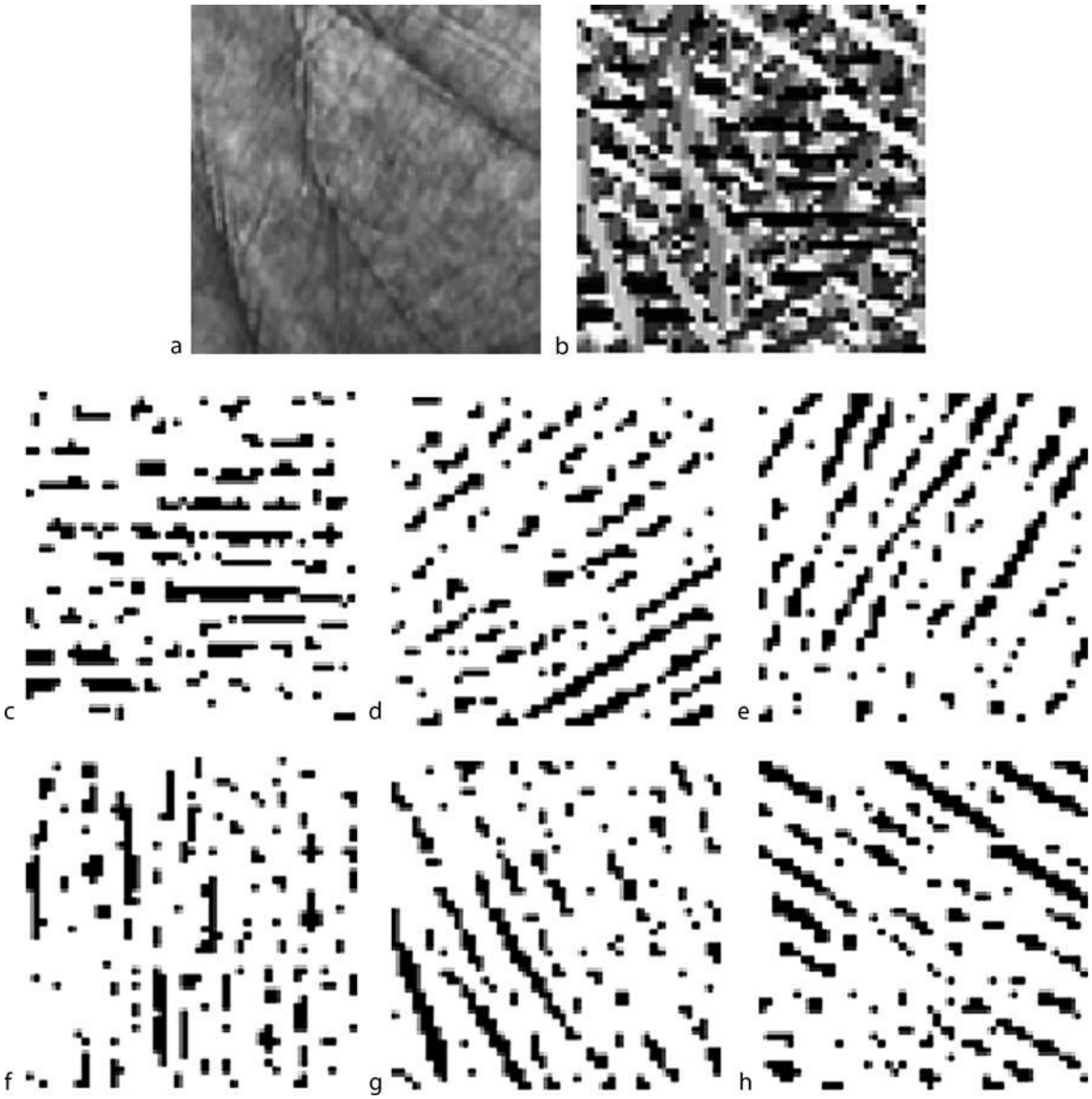
Palmprint Features. **Figure 3** Block Diagram of Fusion Code-based Palmprint Identification.

than 3 to guarantee that the region is shaped like a line. The orthogonal line ordinal filter is defined as the comparison/difference of two orthogonal line-like palmprint image regions. For each local region, three oriented ordinal filters with 0, $\pi/6$ and $\pi/3$ are performed on it to obtain three bit ordinal codes based on the sign of filtering results. Finally, three ordinal templates called the Ordinal Code are obtained as the feature of the input palmprint image.

Line Features

In palmprint images, lines and textures are the most observable features and lines are more appealing than texture to the human eye. The line feature of palmprints can be generalized to principal lines, wrinkles, and creases. Normally, there are three principal lines in a palmprint: the heart line, the head line, and the life line. These lines vary little over time, and their shapes and locations on the palm are the most important physiological features for individual recognition. Most wrinkles are thinner and more irregular than the principal lines; some wrinkles are not only as strong as the principal lines, but are also stable and reliable for identification. Creases cover the entire palm just like ridges in a fingerprint and cannot be observed in low-resolution images.

The principal lines and wrinkles, also called palm lines, are stable and reliable for individual identification and can be exploited and derived from a low-resolution palmprint image. Many algorithms have been developed to extract palm line features for personal authentication. Regarding palm lines as

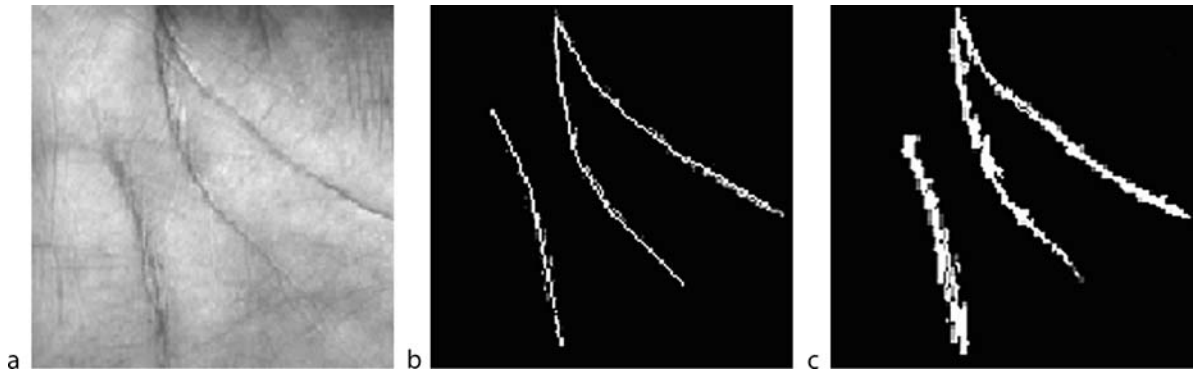


Palmprint Features. Figure 4 An Example of Competitive Code. (**a**) Pre-processed Image. (**c**)–(**h**) The Winning Code 0, 1, 2, 3, 4 and 5, respectively. (**b**) Combination of (**c**)–(**h**).

several straight line segments, line-like features can be extracted by employing some directional masks, as well as Sobel and morphological operations [1, 14]. Applying a wavelet transform followed by directional context modeling, a set of statistical signatures of palm lines can be obtained to characterize palmprints [4].

The structural features of palm lines are a natural choice for palmprint recognition in that they can describe a palmprint clearly and are robust against illumination and noises. According to the properties of palm lines, a set of directional line detectors are devised to detect palm lines in different directions [5]. To ensure that the details of the palm line structure are not lost, the detected lines are finally represented by a chain code which is a pixel-by-pixel direction code of a line.

Except the structural features, the width feature generally reflects the strength information of palm lines (Fig. 5) and is also important for describing the characteristics of a palmprint especially when different palmprints have similar palm line structures. The



Palmprint Features. Figure 5 Structural Feature vs. Width Feature. (**a**) Pre-processed Palmprint Image. (**b**) Its Structural Features. (**c**) Width Features.

width features of palm lines as well as the structural features can be extracted by using a wide line detector [6]. This detector can extract a line completely by inverting the kernel mass which is obtained by an isotropic nonlinear filter.

Appearance Features

In palmprint recognition, the performance of subspace learning methods has been promising [15–17]. These methods efficiently characterize the overall space of raw images by a low-dimensional subspace where standard statistical methods can be used to determine the range of appearance of palmprints.

The Principal Component Analysis (PCA) has been widely used in face recognition and it offers good characterization for palmprint recognition. Based on the Karhunen-Loeve (K-L) transform which is an optimal transform for eliminating statistical correlation, the original palmprint images used in training are transformed into a small set of characteristic feature images, called "eigenpalms," which are the eigenvectors of the training set. Feature extraction is then performed by projecting a new palmprint image into the subspace spanned by the "eigenpalms."

Features of palmprints extracted by the PCA are actually "global" features of all palmprint images implying that they are not necessarily good discriminative representations. The LDA (linear discriminant analysis), based on linear projections, seeks a linear transformation by maximizing between-class variance and minimizing within-class variance and thereby has strong discriminability. Applying the LDA for palmprint recognition, palmprints can be projected from a high-dimensional space to a significantly lower dimensional feature space spanned by Fisherpalms. In this low-dimensional feature space, palmprints from different palms can be discriminated more efficiently.

Both the PCA and LDA methods attempt to find the holistic features of the whole enrollment palmprints and consequently miss the crucial details. Applying the Locality Preserving Projection (LPP) to palmprint images, "Laplacianpalms" are obtained by finding the optimal linear approximation to the eigenfunctions of the Laplacian Beltrami operator on the manifold [17]. They are linear projective maps that arise by solving a variational problem that optimally preserves the neighborhood structure of the dataset. While the Eigenpalm method aims to preserve the global structure of the palmprint image space and the Fisherpalm method preserves the global discriminating information, the Laplacianpalm method strives to preserve the local structure of the palmprint image space.

Summary

Several palmprint features have been investigated for personal authentication and have yielded promising results. According to the performance analysis of various palmprint features presented in the literature, texture-based features, such as ordinal features and competitive coding, revealed the best performance in a large palmprint population. The study of how to effectively combine different palmprint

representations to achieve higher performance has attracted the attention of an increasing number of researchers.

Related Entries

- ▶ Anatomy of Hand
- ▶ Feature Extraction
- ▶ Local Feature Filters
- ▶ Palmprint Matching
- ▶ Palmprint, 3D

References

1. Zhang, D., Shu, W.: Two novel characteristics in palmprint verification: Datum point invariance and line feature matching. Pattern Recognit. **32**, 691–702 (1999)
2. You, J., Li, W.X., Zhang, D.: Hierarchical palmprint identification via multiple feature extraction. Pattern Recognit. **35**, 847–859 (2002)
3. Chen, J., Zhang, C., Rong, G.: Palmprint recognition using crease. In: Proceedings of the 7th International Conference on Image Processing (ICIP), Thessaloniki, Greece, 234–237 (2001)
4. Zhang, L., Zhang, D.: Characterization of palmprints by wavelet signatures via directional context modeling. IEEE Trans. Syst. Man Cybern. Part B. **34**, 1335–1347 (2004)
5. Wu, X.Q., Wang, K.Q., Zhang, D.: Palm-line extraction and matching for personal authentication. IEEE Trans. Syst. Man Cybern. Part A. **36**, 978–987 (2006)
6. Liu, L., Zhang, D.: A Novel palm-line detector. In: Proceedings of 5th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), New York, 563–571 (2005)
7. Zhang, D., Kong, W.K., You, J., Wong, M.: Online palmprint identification. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 1041–1050 (2003)
8. Kumar, A., Zhang, D.: Personal authentication using multiple palmprint representation, Pattern Recognit. **38**, 1695–1704 (2005)
9. Li, W., Zhang, D., Xu, Z.: Palmprint identification by fourier transform. Intern. J. Pattern Recognit. Artif. Intell. **16**, 417–432 (2002)
10. Kumar, A., Zhang, D.: Personal recognition using hand-shape and texture. IEEE Trans. Image Processing **15**, 2454–2461 (2006)
11. Kong, A., Zhang, D., Kamel, M.: Palmprint identification using feature-level fusion. Pattern Recognit. **39**, 478–487 (2006)
12. Kong, A.W.K., Zhang, D.: Competitive coding scheme for palmprint verification. In: Proceedings of 17th International Conference of Pattern Recognition (ICPR), Cambridge, UK, 520–523 (2004)
13. Sun, Z.N., Tan, T.N., Wang Y.H., Li, S.Z.: Ordinal palmprint representation for personal identification. In: Proceedings of IEEE Computer Society International Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, USA, 279–284 (2005)
14. Han, C.C., Cheng, H.L., Fan, K.C., Lin, C.L.: Personal authentication using palmprint features. Pattern Recognit. **36**, 371–381 (2003)
15. Lu, G., Zhang, D., Wang, K.Q.: Palmprint recognition using eigenpalm-like features. Pattern Recognit. Lett. **24**, 1473–1477 (2003)
16. Wu, X.Q., Zhang, D., Wang, K.Q.: Fisherpalms based palmprint recognition. Pattern Recognit. Lett. **24**, 2829–2838 (2003)
17. Wang, J.G., Yau, W.Y., Suwandy, A., Sung, E.: Person recognition by fusing palmprint and palm vein images based on "Laplacianpalm" representation. Pattern Recognit. **41**, 1514–1527 (2008)

Palmprint Matching

Andrew Beng Jin Teoh
Biometrics Engineering Research Center (BERC), School of Electrical and Electronic Engineering, Yonsei University, Seoul, South Korea

Synonyms

Comparison; Dissimilarity; Similarity

Definition

Palmprint matching is a comparison process of two given palmprint and returns either a dichotomy decision (yes or no) or a degree of similarity. Due to the rich features in a palm, including geometrical features (e.g., width, length, area etc. of a palm), principle lines, ridges, singular points, minutiae points, and texture, the matching algorithms require an intermediate palmprint representation to be extracted through a ▶ feature extraction stage. Based on these palmprint features, several approaches to palmprint matching have been devised and they can be broadly classified into two major categories: geometry-based matching and feature-based matching. The integration of two approaches can be done in hierarchical manner to improve the palmprint recognition systems in terms of performance and speed.

Introduction

In general, a palmprint recognition system consists of three components: (1) preprocessing, (2) feature extraction, and (3) matching. A palmprint image will be first preprocessed to determine the region of interest. This process includes segmentation and normalization so that a canonical palmprint image can be produced for sequel processing. Then, an intermediate palmprint representation has to be derived through a feature extraction stage. The common palmprint representation techniques can be mainly divided into five categories according to the palmprint features, such as (1) line-based approaches, (2) texture-based, (3) appearance-based approaches, (4) multiples feature approaches, and (5) orientations-based approaches [1]. A decision is sought eventually through the matching process. The matching is a task to calculate the degree of similarity of two given palmprint and return a decision.

The establishment of palmprint matching strategy is not a trivial problem. The primary reason for studying of palmprint matching is due to the large deviation in different impressions of the same palmprint, i.e., large intra-class variations. The possible factors that contributed to intra-class variations are as follows:

1. Noise. It is introduced by the palmprint acquisition system, e.g., residual prints left on palmprint acquisition device
2. Adverse environmental condition such as bad illumination condition
3. Incorrect interaction by the user with the palmprint recognition system
4. Preprocessing and feature extraction errors. The preprocessing and feature extraction algorithm are imperfect and often introduce measurement errors. The errors will be propagated from the preprocessing stage (e.g., incorrect segmentation and normalization of region of interest) to feature extraction. Another instance is that in low-quality palmprint images, the line-based feature extraction process may introduce a large number of spurious lines or miss to detect the true palm lines

On the other hand, palmprint images from different palms may appear quite similar, i.e., small inter-class variations, especially in palmprint principle lines. A well designed matcher attempts to find the "best fit" between the two palmprint representations, and thus reduces the errors that are introduced by the above sources.

Furthermore, palmprint matching and feature extraction are usually related to each other for both verification and identification problems. Identification problem treats the searching for an input palmprint in a database of M palmprints, thus it can be implemented as a series of consecutive execution of M one to one matches (verification) pairs of palmprints.

Several automatic palmprint matching algorithms have been proposed in the literature of biometrics. Based on the aforementioned feature representations of palmprint image, the palmprint matching approaches can be broadly classified into two categories:

- *Geometry-based matching*: It is a natural way to represent the feature of palmprint using geometry objects, such as points, orientation, lines etc. Lines in palmprint such as principle lines (e.g., head line, life line, and hearth line) and coarse wrinkle are the basic feature of palmprint [2]. A set of feature points or lines segments along the basic palm lines and/or the associated orientations can be extracted from a palmprint image. Geometry-based matching consists of finding the geometrical alignment between the template and the input feature set that returns in the maximum number of features pairing or smallest/biggest degree of ▶ similarity/dissimilarity. However, line features is more popular and widely used as it is relatively easier to be extracted compared to point features in palmprint, even in low spatial resolution palmprint image.
- *Feature-based matching*: Geometry-based matching approach relies heavily on the extraction of points or lines feature from a palmprint image, this might be difficult in very low-quality palmprint images. Alternatively, other features of the palmprint image such as magnitude and orientation information in palm lines can be utilized. The magnitude and orientation of palm lines or texture, in general, can be modeled and extracted by using appearance-based [3, 4], transform-based [5], texture-based [6, 7], and orientation-based techniques [8, 9]. The approaches belonging to this category compare palmprint features based on the similarity/dissimilarity measurement, such as Euclidean distance, hamming distance, angular distance etc. between the two corresponding feature vectors/matrices.

Besides the above two major categories, some other techniques have also been proposed in the literature. For instance, ▶ machine-learning based recognition techniques such as Neural Networks [4], Support Vector Machine [10], and Correlation-based matching [11]. In principle, they could be regarded as sub-categories of feature-based matching according to the feature used, but it is more appropriate to categorize them separately based on the matching approach.

Geometry-Based Matching

Point-Based

In view of structural similarity of fingerprint and palmprint impressions, where both are composed by ridges, it is straightforward to adopt fingerprint minutiae motivated point matching approach to palmprint image. A representative example of point-based matching was proposed by Duta et al. [12]. In their method, given a palmprint image, an average filter is first used to smooth the image followed by binarization process based on a chosen threshold value, *t*. All pixels whose value greater than *t* are labeled as palm line pixels while others are regarded as background pixels. A set of consecutive morphological erosions, dilations, and subtractions are performed to eliminate the spurious palm lines. The outstanding foregrounds pixel locations are subsampled to retrieve a set of 200–400 pixel locations that will be considered as feature points.

For each feature point, the orientation of its corresponding palm line is calculated. The two sets of feature points/orientations are geometrically aligned. The goal of alignment of two feature point sets, *A* and *B* is that to determine a corresponding feature points in *A* such that the highest degree of correlation can be found between *A* and *B* with respect to an optimal transformation, *T*. The matching is finally carried out based on a matching score that is defined as a tuple (P, D), where *P* is the percentage of point correspondences with respect to the minimum number of feature points in set *A* and *B*, and *D* is the average distance between the corresponding points. This matching score was devised based on two sources of intra-class variations, i.e., noise introduced by feature point extraction and non-linear palm deformation due to various finger positions [12], which are modeled by *P* and *D*, respectively. It was shown that in a small size inked palmprint database with 30 subjects, 95% of recognition rate was attained.

Another work on feature point matching was reported by Jane et al. [13]. This work applied an *interesting point* detector, namely Plessey operator to extract the feature points. The interesting point detector differs from the conventional edge detector in the sense that the points detection is based on the how interesting a point is. "Interesting" here refers to a set of application-specific specifications that enables the operator to extract only those representative and distinctive feature points for matching purpose. In general, interesting point detectors will be operated based on a three-step procedure. The first step is to determine a pre-specified size window, based on the average gradient magnitude. This is followed by the classification that distinguishes the types of singular points such as corners, rings, spirals etc. based on a statistical test. The last step is to refine the located point within the window.

For matching purpose, Hausdorff distance is adopted to calculate the degree of similarity of two feature points sets. The Hausdorff distance is a non-linear operator, which measures the degree of the mismatch between two feature point sets *A* and *B*. Mathematically, the Hausdorff distance is $d_H = \max\{h(A, B), h(B, A)\}$ where $h(A, B) = \max_{a\in A}\{\min_{b\in B}\{d(a_i, b_i)\}\}$ and $h(A, B) \neq h(B, A)$. a_i and b_i are feature points of set *A* and *B*, respectively and $d(a_i, b_i)$ is an arbitrary metric between these points. The major advantage of using the Hausdorff distance for matching is that the computation can be accelerated by partitioning the feature point sets *A* into several subsets and match the *B* on these subsets simultaneously. To be specific, the feature point sets can be represented in binary matrices, $A(i, j)$ and $B(i, j)$. The (i, j)th entry indicates an interesting feature point position, which is set to 1 and 0 otherwise. Two distance matrices, D_A and D_B are defined as the distance of each (i, j) location entry to the nearest non-zero location entry of *A* and *B*, respectively. Therefore, the Hausdorff distance, as a function of translation can be computed by considering the point-wise maximum of all the translated D_A and D_B in the form of $d_H(i, j) = \max(\alpha, \beta)$ where $\alpha = \max_a D_A(a_i - i, a_j - j)$ and $\beta = \max_b D_B(b_i + i, b_j + j)$ [13].

A limitation of the feature points matching techniques, which based on exhaustive scanning approach is cumbersome and may not meet the real-time requirement for on-line matching in a large database.

Line-Based

Compared to the point-based matching, line-based matching is conceived more informative than the point-based matching in palmprint recognition system due to the rich line features in a palmprint image. Line-base matching technique first extract line feature, which is composed by curves and straight line. In [2], edge filters are applied to extract principle lines, thick wrinkles, and ridges at various orientations repeatedly and combine them with a post-processing algorithm by line linking and thinning at the final stage. The representation of each extracted line segment is determined by a series of end points: $(u_1(i), v_1(i))$ and $(u_2(i), v_2(i))$, $i = 1, \ldots, m$ where m is the number of line segments. The end points coordinate is described by a two-dimensional right coordinates system, which is uniquely determined by the datum points. In general, each line segment can be categorized by three parameters: (1) slope, m, (2) intercept, c and (3) angle of inclination, α. They can be calculated using equations $m(i) = (v_2(i) - v_1(i))/(v_2(i) - v_1(i))$, $c(i) = v_1(i) - u_1(i)m(i)$, $\alpha(i) = \tan^{-1}(m(i))$. The Euclidean distances between the endpoints of two line segments, d_1 and d_2 is given as $d_1 = \sqrt{(u_1(i) - u_1(j))^2 + (v_1(i) - v_1(j))^2}$ and $d_2 = \sqrt{(u_2(i) - u_2(j))^2 + (v_2(i) - v_2(j))^2}$.

For matching purpose, several conditions were proposed as follows: (1) If both d_1 and d_2 are less than a predefined threshold value, t, then this implies that both line segments are the same. (2) If the difference between α and c is less than sum of their respective threshold values, then this implies two line segments have the equal α and c. Among the class of equal α and c, if one of d_1 and d_2 is less than t, then two line segments are considered identical. (3) When two line segments overlap, they are regarded as a single line segment if the midpoint of one line segment is between two endpoints of another line. Based on the three rules given, a corresponding pair of palm lines can be obtained. A decision criterion is hence defined as $r = 2N/(N_1 + N_2)$, $0 < r < 1$ where N is the number of these corresponding pairs; N_1 and N_2 are the numbers of the line segments determined from two palmprint images, respectively. In a medium scale inked palmprint database that consists of 200 subjects, it was reported that 92% of recognition rate can be achieved.

Geometry-based matching approach, especially line features is an active research area and still evolving. The researchers believe that line-based features in palmprint are highly discriminant. Huang et al. evidences that even simple line-based features such as principle lines also can show the high discriminability [1]. In this work, they proposed to use a modified finite Radon transform to extract the principle lines and represent them in a binary matrix with size $h \times k$, where principle line point is set to 1 and 0 for others. A new matching strategy, known as pixel-to-area comparison was devised based on the pixel to area comparisons for robust line matching. Given two principle line matrices, A and B, the matching score from A to B is defined as

$$S = \max(s(A, B), s(B, A)), 0 \leq S \leq 1$$

where

$$s(A, B) = \left(\sum_{i=1}^{h} \sum_{j=1}^{k} A(i, j) \cap \overline{B}(i, j) \right) / N_A \text{ and}$$

$$s(B, A) = \left(\sum_{i=1}^{h} \sum_{j=1}^{k} B(i, j) \cap \overline{A}(i, j) \right) / N_B.$$

Here, $\cap$ denotes a logical "AND" operation, N_A and N_B are the number of points on detected principle lines in A and B. $\overline{B}(i, j)$is a small area around $B(i, j)$ and is defined as $B(i + 1, j)$, $B(i-1, j)$, $B(i, j)$, $B(i, j + 1)$ and $B(i, j-1)$. The same definition applies to $\overline{A}(i, j)$. S is devised in such a way that it is robust to slight translations and rotations between the two images, with limited to one pixel translation and 3° rotation. In practice, the translation might be large due to imperfect preprocessing. This problem can be alleviated by translating one image vertically and horizontally repeatedly in the range of -2 to 2 pixels and match with another image. The maximum value of S is regarded as a final decision score. The experiment result showed that the method could achieve an equal error rate (EER) of 0.565% in a large scale database

with 386 palmprints. It is believed that the performance can be improved further if other line features such as wrinkles are included.

Feature-Based Matching

Feature-based matching utilizes magnitude and orientation information of palm lines, or texture in general for matching purpose. Palm lines magnitude information can be modeled and extracted by using statistical and algebraic techniques such as appearance-based feature representation [3, 4] (e.g., principle component analysis (PCA), fisher discriminant analysis (FDA), Independent component analysis (ICA) etc), Fourier spectrum [5], wavelet transform [10], discrete cosine transform [14], and convolution masks [13]. On the other hand, orientation information of palm lines can be effectively extracted by using Gabor filters [6, 8] and ordinal representation [9].

The common feature representation based on statistical and algebraic techniques usually appear in either one-dimensional feature vector with length n, $\boldsymbol{v} = \{v_i | \ i = 1, \ldots, n\}$ or two-dimensional feature matrix with size $m \times n$, $\boldsymbol{V} = \{V_{ij} | \ i = 1, \ldots, m, \ i = 1, \ldots, n\}$. For instance, [14] characterize a palmprint image by using a set of context based wavelet signatures. Specifically, a palmprint image is decomposed into J wavelet scales and only three detail wavelet sub-band coefficients, i.e., horizontal, vertical, and diagonal. For each sub-band coefficient, four statistical readings, namely (1) the Average Gravity Center Signature (AGCS), (2) the Density Signature (DS), (3) the Spatial Dispersivity Signature (SDS), and (4) the Energy Signature (ES) can be computed. Each palmprint image will generate a feature vector with length $k = 3 + 3*3*J$, which consists of three pairs of AGCS, $3J$ DS, $3J$ SDS and $3J$ ES.

On the other hand, algebraic techniques such as PCA, FDA, ICA etc first transform the palmprint training images into a small set of characteristic feature images, called Projection Matrix, which are the eigenvectors of the training set. Then, feature extraction is performed by projecting a new palmprint image with length n into the subspace spanned by the Projection Matrix. The output is a feature vector with length $k << n$.

For the representation in one-dimensional feature vector, the matching is normally done by using distance metric such as city block distance, $d_1 = |\boldsymbol{v}_i - \boldsymbol{v}_j|$, Euclidean distance, $d_2 = \sqrt{\sum \left(\boldsymbol{v}_i - \boldsymbol{v}_j\right)^2}$, weighted Euclidean distance, $d_2 = \sqrt{\sum_k \frac{1}{\mathbf{w}_k}\left(\boldsymbol{v}_i - \boldsymbol{v}_j\right)^2}$, angular distance, $d_3 = \frac{\boldsymbol{v}_i{}^T \boldsymbol{v}_i}{\|\boldsymbol{v}_i\| \|\boldsymbol{v}_j\|}$ etc., where $i \neq j$ and $\boldsymbol{w}$ is a weight vector. Without loss of generality, the degree of similarity/dissimilarity of these distance metrics is given in term of score between 0 and 1.

Another popular palmprint feature extractor used in extracting the texture information is 2-D Gabor filter and its variants [6–8]. In this technique, a 2-D Gabor filter will convolute with a preprocessed palmprint image followed by a robust encoding to convert the convoluted output into two binary matrices, which corresponding to real and imaginary parts of the output. Given two palmprint feature binary matrices, A_R (A_I) and B_R (B_I) with size $h \times h$, the matching is performed via a normalized hamming distance H_1 such as

$$H_1 = \frac{\sum_i \sum_j (A_R(i,j) \otimes B_R(i,j) + A_1(i,j) \otimes B_1(i,j))}{2h^2}$$

where $\otimes$ is an bitwise Ex-OR operator [6]. The equation can be modified to cattle the translation problem with

$$H_2 = \min_{|s|<S, |t|<T} \left\{ \left[\sum_{i=max(1,1+s)}^{min(h,h+s)} \sum_{i=max(1,1+t)}^{min(h,h+t)} A_R(i+s, j+t) \otimes B_R(i,j) + A_I(i+s, j+t) \otimes B_I(i,j) \right] \Big/ 2H(s)H(t) \right\}$$

s and t are set to 2 based on the assumption that translation (both vertically and horizontally) is limited in the range of 2 pixels and $H(s) = \min(h, h+s) - \max(1, 1+s)$. However, this metric does not consider rotation invariant, but this issue can be alleviated during the enrolment stage. For instant, rotate the coordinate system by a few degrees and perform feature extraction [6].

If a palmprint is not segmented properly during the preprocessing stage, a number of non-palmprint pixels will be introduced in the extracted feature matrix. In this circumstance, these pixels are detected by using some simple thresholding methods and their locations can be recorded in the mask matrices, A_M and B_M that corresponded to the feature matrices A and B. H_1 can

be modified to

$$H_3 = \frac{\sum\sum A_M \cap B_M \cap (A_R \otimes B_R) + A_M \cap B_M \cap (A_I \otimes B_I)}{2\sum\sum A_M \cap B_M.} \quad [7].$$

On the other hand, [8] perceived that orientation information from the palm lines could be the prominent features for a palmprint image. An even Gabor filter with six different orientations is used to convolute with the palmprint image and their contrast magnitudes are sought. Based on the winner-take-all competitive principle, the index (ranging from 0 to 5) of the minimum contrast magnitude is represented by three bits, namely competitive code. The matching of two inputs, A and B can be carried out through

$$H_4 = \frac{\sum\sum\sum_{k=0}^{3} A_M \cap B_M \cap \left(A_k^b \otimes B_k^b\right)}{3\sum\sum A_M \cap B_M},$$

where $A_k^b(B_k^b)$is the ith bit plane of $A(B)$.

As far as the performance is concerned, orientation-based feature representation couples with matching in hamming domain are deemed to be the most promising techniques in palmprint recognition. In a comparison study done in PolyU (medium size) and UST datasets (large size), the competitive code showed better performance with EER = 0% and EER = 0.38%, respectively compare to PalmCode (0.34%, 1.68%) and FusionCode (0.11%, 0.75%). However, the Ordinal code, which is another orientation-based feature representation combined with robust encoding show the best performance in terms of EER = 0% and EER = 0.22% [9]. In addition to that, low computation complexity and small template representation are also another compelling advantage of this approach.

Summary

The establishment of palmprint matching methodology is highly related to feature representation of palmprint image and it is essential as the preprocessing and feature extraction are imperfect. Geometry-based matching approach was first proposed in view of natural representation of palmprint in feature points and lines basis. However, it was believed that they are less reliable as features point and lines are difficult to be explicitly extracted, but recently it is receiving renewed interest. Feature-based matching approach, as an alternative, which utilizes the texture information (magnitude and orientation of palm lines) of palmprint, has also shown the significant advantages in terms of performance, representation compactness, and low computation complexity. The integration of two approaches, e.g., in hierarchical manner [15] could be the promising way to improve the palmprint recognition systems in performance and speed.

Related Entries

- ▶ Authentication
- ▶ Identification
- ▶ Verification

References

1. Huang, D., Jia, W., Zhang, D.: Palmprint verification based on principal lines. Pattern Recognit. **41**(4), 1316–1328 (2008)
2. Zhang, D., Shu, W.: Two novel characteristics in palmprint verification: datum point invariance and line feature matching. Pattern Recognit. **32**(4), 691–702 (1999)
3. Tee, C., Teoh, A.B.J., Goh, M.K.O., Ngo, D.C.L.: An automated palmprint recognition system. Image Vis. Comput. **23**(5), 501–515 (2005)
4. Shang, L., Huang, D.S., Du, J.S., Zheng, C.N.: Palmprint recognition using FastICA algorithm and radial basis probabilistic neural network. Neurocomputing **69**, 13–15 (2006)
5. Li, W.X., Zhang, D., Xu, Z.Q.: Palmprint Identification by Fourier Transform. Int. J Pattern Recognit. Artif. Intell. **16**(4), 417–432 (2002)
6. Kong, W., Zhang, D., Li, W.: Palmprint feature extraction using 2-D Gabor filters. Pattern Recognit. **36**(10), 2339–2347 (2003)
7. Zhang, D., Kong, W., You, J., Wong, M.: Online Palmprint Identification. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1041–1050 (2003)
8. Kong, W., Zhang, D.: Competitive coding scheme for palmprint verification. In: Proceedings of the 17th International Conference on Pattern Recognition. **1**, 520–523 (2004)
9. Sun, Z.N., Tan, T.N., Wang, Y.H., Li, S.Z.: Ordinal palmprint representation for personal identification. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. 279–284 (2005)

10. Zhou, X.H., Peng, Y.H., Yang, M.: Palmprint Recognition Using Wavelet and Support Vector Machines. Lect. Notes Comput. Sci. **4099**, 385–393 (2006)
11. Hennings-Yeomans, P.H., Kumar, B.V.K.V., Savvides, M.: Palmprint Classification Using Multiple Advanced Correlation Filters and Palm-Specific Segmentation. IEEE Trans Inf Forensic Security **2**(3), 613–622 (2007).
12. Duta, N., Jain, A.K., Mardia, K.V.: Matching of palmprints. Pattern Recognit. Lett. **23**(4), 477–485 (2002)
13. You, J., Li, W., Zhang, D.: Hierarchical palmprint identification via multiple feature extraction. Pattern Recognit. **35**(4), 847–859 (2002)
14. Kumar, A., Zhang, D.: Personal recognition using hand-shape and texture. IEEE Trans. Image Process. **15**(8), 2454–2461 (2004)
15. You, J., Kong, W., Zhang, D., King Hong Cheung.: On hierarchical palmprint coding with multiple features for personal identification in large databases. IEEE Trans. Circ. Syst. Video Tech. **14**(2), 234–243 (2004)

Palmprint Recognition, 3D

▶ Palmprint, 3D

Palmprint Representation

▶ Palmprint Features

Palmprint Sensor

▶ Fingerprint, Palmprint, Handprint and Soleprint Sensor

Parallel Fusion Network

A parallel fusion network is a fusion network topology in which the sensor nodes are connected directly to the central fusion processor. There is no distribution of the processing through the network.

▶ Fusion, Decision-Level

Parametric Models

Synonym

Mathematical models

Definition

Parametric or mathematical models are compact math and algorithmic representations that use variables that when defined allow the creation of the modeled phenomenon that closely approximate the actual phenomenon that is being modeled. An example of an empirically derived parametric model is the University of Bologna's SFinGe fingerprint generator.

▶ Biometric Sample Synthesis

Parametric-Based Biometrics

▶ Biometric Sample Synthesis

Part-Based Face Recognition

▶ Face Recognition, Component-Based

Partial Occlusion

Occlusions of a local region of the face with objects such as sunglasses, scarf, hands, and hair are generally

called partial occlusions. Partial occlusions can, in theory, correspond to any occluding object. Generally, the occlusion has to be of less than 50% of the face to be considered a partial occlusion.

► Face Recognition, Component-Based

Passive Biometrics

In passive biometrics systems the subject does not have to take active part in the process of identification/verification or, in fact, does not even know that the process of identification takes place. In such biometrics the user does not have to cooperate with the system and does not need to touch any device or perform any action.

► Ear Biometrics

Patron Format Specification

► Common Biometric Exchange Formats Framework Standardization

Pattern Recognition

Pattern recognition aims to classify data (patterns) based on either a prior knowledge or on statistical information extracted from the patterns. The patterns to be classified are usually groups "i" of measurements or observations, defining points in an appropriate "i" multidimensional space. A complete pattern recognition system consists of a sensor that gathers the observations to be classified or described; a feature extraction mechanism that computes numeric or symbolic information from the observations; and a classification or description scheme that does the actual job of classifying or describing observations, relying on the extracted features.

► Universal Background Models

PCA (Principal Component Analysis)

► Deformable Models
► Face Alignment
► Face, Forensic Evidence of
► Hand Shape
► Linear Dimension Reduction
► SFinGe
► Soft Biometrics

Pedestrian Detection

► Human Detection and Tracking

Peg

Short projecting pin used for marking position. Within Hand Geometry biometrics, they may be placed on a planar surface, to guide the user when placing his or her hand.

► Hand Geometry

Pen Altitude

Angle measured counter-clockwise from the perpendicular projection of the pen onto the writing plane to the pen. Altitude values can be acquired

as time-series data. Altitude is a factor representing the pen inclination, and it is used together with azimuth. These features are used for on-line signature verification.

► Signature Recognition

Pen Azimuth

Angle measured clockwise from the positive y axis to the perpendicular projection of the pen onto the writing plane. Azimuth values can be acquired as time-series data. Azimuth is a factor representing the pen inclination, and it is used together with altitude. These features are used for on-line signature verification.

► Signature Recognition

Pen Inclination (Pen Tilt)

Angle characterizing how the pen is held.

How the pen is held depends on the person, and it changes during the signing process. Thus, this feature is useful for on-line signature verification. Pen inclination has two degrees of freedom and it can be represented in two ways. One way is to represent it by angles measured from two axes. The other way is to represent it by azimuth and altitude.

► Signature Recognition

Pen Pressure

Force of the pen on the writing surface.

Pen pressure values can be acquired as time-series data. This feature is used for on-line signature verification. Generally, pen pressure is measured with a stylus pen and tablet, or a pressure sensitive pen. Pen pressure is represented by N quantized levels, for example 1024 or 512. If the pen pressure is 0, the pen does not touch the writing surface. Some kinds of tablet can detect the pen position even when the pen does not touch the tablet.

► Signature Recognition

Pen Tablet

A digitizing or pen tablet is a flat device that allows recording handwriting movements. Usually these devices are based on an electromagnetic principle. The tablet has an embedded wire grid which acts as a transmitter. The pen (which is specifically designed for the tablet) acts as an antenna, which resonates and emits a signal that is captured by the tablet, allowing to detect its position with high accuracy. This allows the tablet to detect the pen movement even if it is not in contact with the tablet (in a reasonable range of proximity).

► Signature Databases and Evaluation

Penetration Rate

The ratio of fingerprints retrieved over the size of the database.

► Fingerprint Indexing

Pen-tip Position (Pen Coordinates)

Position of the pen-tip on the writing plane.

The pen-tip position is generally represented using Cartesian coordinates $(x; y)$. The time series of

x-coordinates $x(t)$ and y-coordinates $y(t)$ are acquired by a capture device at a sequence of sample points. $x(t)$ represents the position on the horizontal axis and $y(t)$ represents the position on the vertical axis at each time t. This feature is used for on-line signature verification. Concatenating consecutive sample points $(x(t); y(t))$ reproduces the shape of the signature.

► Signature Recognition

Perceptual Expertise

Perceptual expertise is the change in perception due to experience that may come about through a variety of mechanisms such as shifting attention to different dimensions or features, or new processes such as configural processing.

► Latent Fingerprint Experts

Performance Bias in Synthesized Biometric Data

Performance bias as applied to synthesized biometric data is an asystematic error experienced by a biometric system. This type of bias can arise from a lack of or an excessive presence of some synthetically generated distortions that appear in a larger amount or do not present at all in real biometric data.

► Iris Sample Synthesis

Performance Evaluation Measures

► Performance Measures

Performance Evaluation, Overview

Shiguang Shan[1], Xilin Chen[1], Wen Gao[1,2]
[1]Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[2]Peking University, Beijing, China

Synonyms

Biometric technology test; Performance testing

Definition

Performance evaluation assesses accuracy and usability of biometric algorithms or systems. Performance measures are computed for verification, identification, and watch list tasks, in order to either discover the state-of-the-art of biometric technologies or quantify how well a biometric system meets the requirements of specific applications. Evaluation protocols and biometric databases for testing should be carefully designed to avoid biased results or conclusions.

Introduction

In the past several decades, many biometric algorithms and commercial biometrics systems have emerged. Many of them have reported very impressive results on some public or private databases in all kinds of publications. However, it is hard to compare them based only on the reported results, due to the difference in either the evaluation methods or the testing databases they exploited. Actually, these published results often lead to confusion in public: application users are puzzled when choosing product venders and researchers (especially green hands) may not be able to clearly know the state of the art. Therefore, it is indeed very important to standardize methodology for performance evaluation of biometric technologies.

For most biometric technologies, there are two main tasks when applying them in practice, i.e.,

verification and identification. The former needs to answer "Is he who he says he is?", while the latter cares about "who is he?". According to whether the unidentified end-user is enrolled in the system, identification is further categorized into two types: ► closed-set identification and ► open-set identification. Typical verification application is access control, while closed-set identification can be applied to mug shot retrieval for instance. In surveillance scenario, open-set identification is also named "watch list," which aims at answering "Is he one of the persons of interest?" generally in real time. For example, face recognition, gait recognition, and speaker recognition can be applied for this purpose in a surveillance scenario since they can work in nonintrusive mode.

Performance evaluations can also be categorized into three different types: algorithm evaluation, scenario evaluation, and operational evaluation, as is described in evaluation protocols part of this essay.

Performance Measures

Evidently, different tasks should explore distinct performance measures. For verification, receiver operating characteristic (ROC) curve is generally used to show the trade-off between two error rates: false reject rate (FRR) versus false accept rate (FAR). Sometimes, the equal error rate (EER) point on the ROC, where FRR is equal to FAR, is used as a single measurement. As for identification, identification rate, rank-*k* identification rate, or cumulative match characteristic (CMC) is often used to compare different techniques. The reader is referred to the ► performance measures entry for more details.

For watch list applications, in some sense, it is the verification of the rank-1 identification. So, its performance can be measured by the identification rate at certain pre-defined FAR, say 0.1%. This measure is exploited by face recognition vender test (FRVT) [1].

Except the accuracy measures mentioned above, there are also some performance criteria measuring the usability of biometric systems, such as ► failure to acquire rate, ► enrollment time, ► response time, ► throughput, and ► scalability. The readers are referred to the definitional entries for their description.

Datasets

The abovementioned performance measures are generally obtained by testing the biometric systems on some databases. Evidently, the performance of a system or algorithm depends not only on its capacity but also the characteristic of the database. So, here it is worth noting that pure recognition accuracy, say 100%, means nothing if the database is not described clearly. The following factors about validation database must be considered carefully when performance evaluation is conducted.

Here, first several distinct datasets in evaluation: training dataset, validation dataset, gallery, and probe dataset need to be distinguished. Among them, the *training set* is used for learning the biometric models and designing the recognition algorithms, including the feature extractor (e.g., principal component analysis and discriminant analysis) and the classifier. *Validation set* is used to tune the parameters of the leaned models or the algorithms, for instance the dimension of the feature vector or some empirical thresholds. In some literature, training set and validation set are combined together and called *training set* commonly. For instance, in FVC2006 [2], a subset of fingerprint impressions acquired with various sensors was provided to registered participants, to allow them to adjust the parameters of their algorithms. The *gallery* here means the dataset containing all the registered biometric traits of all the enrolled users in the system, that is, the templates for each enrolled users are extracted from this dataset. Note that, the gallery is often taken as part of (or the same as) the training set by many researchers. This is mostly acceptable; however, in many applications, each enrolled subject may register only one biometric sample, which implies that it is impossible to train a feature extractor (e.g., Fisher discriminant analysis) or classifier. In this case, a training set is necessary. The *probe* dataset contains testing biometric samples that need to be recognized by matching against the templates in the gallery. Note that, for identification task, all the subjects in the probe set can be registered subjects (i.e., with at least one template), while for verification and watch list applications, part of the subjects in the probe set should be unregistered subjects, which is used as impostors to estimate the false accept rate. In literature, the gallery and probe set are called together

the *testing set*. Sometimes, gallery is also called *target set*, while probe set is also called *query set*, e.g., in FRGC [3].

When collecting biometric samples in the abovementioned four datasets, some points should be considered carefully. First, the samples in the gallery should never be included in the probe set also, since this will definitely result in correct match. Secondly, whether the samples in the testing set should be contained in the training set is task-dependent. Thirdly, whether the subjects in the training and testing set are overlapped partially or completely is application-dependent. For instance, in face recognition technology (FERET) evaluation [4], part of the face images (and subject) in the gallery and probe sets are also in the training set for algorithm development. However, in FRVT [1], all the images (and subjects) in the testing set are confidential to all the participants, which means the developers have to consider carefully how their algorithms can generalize to unseen subjects. In contrast, the Lausanne Protocol based on XM2VTS database [5] does not distinguish the training set from the gallery, i.e., the gallery is the same as the training set. Evidently, different protocol will result in evaluation of different difficulty.

Difficulty Control

The goals of performance evaluation are multifold, such as to compare several algorithms and choose the best one, or determine whether one technology can meet the requirements of specific applications. So, it is very important to control the difficulty of the evaluation. The evaluation itself should not be too hard or too easy. If the evaluation is too easy, all the technologies might have similarly perfect performances and thus no statistically salient difference can be observed among the results. Similarly, if the evaluation is too challenging, all the systems may not work and have bad performance. Therefore, it is indeed very important to control the difficulty of the evaluation in order to make the participants perform discrepantly.

The difficulty of an evaluation protocol is mainly determined by the *variations* of the samples in the probe set from the registered ones. The more the variation, the more difficult the evaluation is. The sources of the variations are multifold. Coarsely, they can be categorized into two classes: *intrinsic variations* and *extrinsic variations*. The former means the changes of the biometric feature itself, while the latter comes from the external factors especially during the sensing procedure. For instance, in face recognition, variations in the facial appearance due to the expression and aging are intrinsic, while those due to lighting, viewpoint, camera difference, and partial occlusion are extrinsic. For instance, more recently, multiple biometric grand challenge (MBGC) [6] is being organized to investigate, test, and improve performance of face and iris recognition technology on both still and video imagery through a series of *challenge* problems, such as low resolution, off-angle images, unconstrained face imaging conditions etc. Especially, for all biometrics, the time interval between the acquisition of the registered sample and unseen samples presented to a system is an important factor, because different acquisition time implies both intrinsic and extrinsic variations. For an evaluation of academic algorithms, a reasonable distribution of all the possible variations in the testing set is desirable, while for application-specific system evaluation it is better to include variations most possibly appearing in the practical applications.

The abovementioned database structure also affects the difficulty of the evaluation. If the samples or the subjects in the testing set have been included in the training set, the evaluation becomes relatively easy. If all the testing samples and subjects are novel to the learned model or system, the overfitting problem might make the task more challenging. In extreme case, if the training set and the testing set are heterogeneous, the task will be much more difficult. For instance, if the training set contains only biometric samples of Mongolian, while the testing samples are from the western. Therefore, the structure of the database for evaluation should be carefully designed to tune the difficulty of the evaluation.

Another factor influencing the evaluation difficulty is the database size, i.e., the number of registered subjects in the database. This is especially important for identification and watch list applications, since evidently the more subjects to recognize, the more challenging the problem becomes. Some observations and conclusions have been drawn in FRVT2002 [1].

Finally, the number of registered biometric samples for each subject should also be considered to control the difficulty of the evaluation. Generally speaking, the task will become easier with the increase of the sample number per person. Note that, an algorithm that works very well with many samples per person does not necessarily work similarly well when the number of samples for each subject is very few. For instance, in many face recognition applications, there might be only one registered face image per person in the database.

Evaluation Protocols

An evaluation protocol determines how to test a system, design the datasets, and measure the performance. Successful evaluations should be administered by third parties. The details of the evaluation procedure must be published along with the evaluation protocol, testing procedures, performance results, and the dataset (at least some representative examples). Also, the information on the evaluation and data should be sufficiently detailed so that users, developers, and vendors can repeat the evaluation [7]. Generally, there are three types of evaluation as described in the following.

Algorithm Evaluation

This kind of evaluation assesses biometric technology itself. Laboratory or prototype algorithms are evaluated to measure the state of the art, to define the technological progress, and to identify the most promising approaches. Typical technology evaluation protocols include the FERET series of face recognition evaluations [4], the National Institute of Standards and Technology (NIST) speaker recognition evaluations [8], the Lausanne Protocol based on XM2VTS database [5], FRGC [3], and the evaluation protocol based on CAS-PEAL-R1 face database [9]. In this kind of evaluation, all the systems are generally tested with completely the same dataset for the purpose of fairness. As mentioned above, some of the protocols provide training set for algorithm development, and in terms of intrinsic and extrinsic variations the training samples are homogeneous with those in the testing set. For this kind of evaluation, accuracy measures are the main performance criteria.

Scenario Evaluation

This type of evaluation type aims at checking whether a biometric technology is sufficiently mature to meet the requirements for a class of applications. In this case, because the systems might have their own data acquisition sensors, the systems are tested with slightly different data [7]. To compensate for this difference, the evaluation must be designed carefully to evaluate systems under as close condition as possible. In addition, since the evaluations are conducted under real-world field conditions, they cannot be repeated exactly. As a kind of system evaluation, performance criteria measuring both accuracy and usability of the system are required to consider, such as failure to acquire rate, enrollment time, response time, throughput, and scalability.

Operational Evaluation

Instead of evaluating for a class of applications, an application-specific evaluation measures the performance of a specific system for a specific application. For example, an application-specific evaluation might need to measure the performance of system X on verifying the identity of people as they enter secure building Y. The primary goal of this kind of evaluation is to determine if a biometric system meets the requirements of a specific application [6]. The performance measures are generally the same as those of scenario evaluation.

Summary

Performance evaluation should tell unbiased facts. To this goal, the independent evaluators must deeply investigate the requirements of the applications concerned, determine the planned difficulty of the evaluation and collect appropriate datasets, then assess the systems with suitable performance measures, and finally report the results along with the evaluation procedure in detail.

Related Entries

- ▶ Evaluation of Gait Recognition
- ▶ Face Databases and Evaluation
- ▶ False Match Rate
- ▶ False Non-Match Rate
- ▶ Fingerprint Databases and Evaluation
- ▶ Hand Databases and Evaluation; Iris Databases
- ▶ Identification
- ▶ Influential Factors to Performance
- ▶ Iris Challenge Evaluation
- ▶ Iris Recognition Performance under Extreme Image Compression
- ▶ Large Scale Evaluation
- ▶ Performance
- ▶ Performance Measures
- ▶ Performance Testing Methodology Standardization
- ▶ Reference Set
- ▶ Speaker Databases and Evaluation
- ▶ Verification

References

1. NIST: Face recognition vendor test (FRVT). http://www.frvt.org, 2000, 2002, 2004
2. The fourth international fingerprint verification competition. http://bias.csr.unibo.it/fvc2006/
3. Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., et al.: Overview of the face recognition grand challenge. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 947–954. IEEE Computer Society, Washington, DC (2005)
4. Phillips, P.J., Moon, H., Rauss, P., Rizvi, S.A.: The FERET evaluation methodology for face-recognition algorithms. In: Proceedings of the IEEE International conference on Computer Vision and Pattern Recognition (CVPR'97), pp.137–143. IEEE Computer Society, Washington, DC (1997)
5. Messer, K., Matas, J., Kittler, J., et al.: XM2VTSDB: the extendedM2VTS database. In: Second International Conference on Audio and Video based Biometric Person Authentication, March 1999, Washington, DC (1999)
6. NIST: Multiple biometrics grand challenge. http://face.nist.gov/mbgc/. Accessed Aug 2008
7. Phillips, P.J., Martin, A., Wilson, C.L., et al.: An introduction to evaluating biometric systems. IEEE magazine on computer, pp. 56–63 (2000)
8. NIST: Spoken language technology evaluations. http://www.nist.gov/speech/tests/sre/index.html. Accessed Aug 2008
9. Gao, W., Cao, B., Shan, S., Chen, X., et al.: The CAS-PEAL large-scale chinese face database and baseline evaluations. IEEE Trans. Syst. Man Cybern. (Part A) **38**(1), 149–161 (2008)

Performance Measures

Jihyeon Jang, Hale Kim
Inha University, Incheon, Korea

Synonyms

Performance evaluation measures; Performance metrics

Definition

Performance measures in biometrics define quantifiable assessments of the processing speed, recognition accuracy, and other functional characteristics of a biometric algorithm or system. The processing speed is evaluated by *Throughput rate* which represents the number of users that can be processed per unit time, and the typical metrics for recognition accuracy are the rates of *Failure-to-enroll*, *Failure-to-acquire*, *False non-match*, *False match*, *False reject*, and *False accept*. In addition to these fundamental performance measures, there are other measures which are specifically dependent on applications (verification, open-set identification, or closed-set identification), such as *False-negative* and *False-positive identification error rates*. Also, graphic measures such as *DET curve*, *ROC curve*, and *CMC curve* are very efficient tools to present overall matching performance of biometric algorithms or systems. Biometric performance testing focuses on the evaluation of technical performance and various error rates of biometric algorithms or systems.

Introduction

The purpose of biometric performance evaluation is to determine the range of errors and throughput rates, with the goal of understanding and predicting real-world recognition and throughput performance of biometric systems. The error rates include both false-positive- and false-negative decisions as well as failure-to-enroll and failure-to-acquire rates across the test population. Throughput rates refer to the number of users processed per unit time, based on both computational speed and human–machine interaction. These measures are defined to be applicable to all biometric systems and devices.

In general, biometric performance testing is divided into three categories: *technology*, *scenario*, and *operational* [1–3]. The following summarizes the characteristics and differences of evaluation types, especially focusing on the resulted metrics.

Technology Evaluation

Technology evaluation is an offline process for testing biometric components using a precollected corpus of samples. Its goal is to compare the performance of biometric algorithms for the same biometric modality. Only algorithms compliant with a given input/output protocol are tested. Although sample data may be distributed for developmental or tuning purposes prior to the test, the actual testing must be done on data that have not been previously seen by algorithm developers. The test results are repeatable because the test corpus is fixed, and provide most of the performance metrics.

Scenario Evaluation

Scenario evaluation is an online process for determining the overall system performance in a prototype or simulated application. Testing is performed on a complete system in an environment that models a real-world target application. Each tested system has its own acquisition devices, while data collection has to be carried out across all tested systems with the same population in the same environment. Test results are repeatable only to the extent to which the test scenario and population can be carefully controlled, and provide only predicted end-to-end throughput rates and error rates.

Operational Evaluation

Operational evaluation is also an online process whose goal is to determine the performance of a complete biometric system in a specific application environment with a specific target population. In general, its test results are not repeatable because of uncontrolled operational environments and population. This evaluation provides only end-to-end throughput rates, false accept, and false reject rates.

This article restricts discussion to the performance measures of technology evaluation because they are mathematically well defined and used more often in real-world biometric performance evaluation.

In a technology evaluation, biometric systems or algorithm components are evaluated with afixed corpus of samples collected under controlled conditions. This allows direct comparison among evaluated systems, assessments of individual systems' strengths and weaknesses, or insight into the overall performance of the evaluated systems. Examples of benchmark test evaluations are Facial Recognition Technology (FERET) [4, 5], Face Recognition Vendor Test (FRVT) 2000, 2002 [6, 7], 2006, Fingerprint Verification Competition (FVC) 2000, 2002, 2004 [8–10], 2006, and National Institute of standards Technolgy (NIST) Speaker Recognition Competitions [11, 12]. Not only the performance metrics but also the test protocols introduced by these technology evaluations have become the basis of the ISO/IEC standards on biometric performance testing and reporting.

The international standards for testing and reporting the performance of biometric systems have been studied and developed by the Working Group 5 of ISO/IEC JTC 1's Subcommittee 37 on Biometrics, one of which is ISO/IEC IS 19795 consisting of the following multiparts described in Table 1, under the general title *Information technology – Biometric performance testing and reporting*. ISO/IEC 19795 is concerned solely with the scientific "technical performance testing" of biometric systems and devices. Especially, ISO/IEC 19795-1 presents the requirements and best scientific practices for conducting technical performance testing. Furthermore, it specifies performance metrics for biometric systems. Most of the

Performance Measures. Table 1 Biometric performance testing and reporting standards by ISO/IEC JTC 1/SC 37

Standard No.	Subtitle
19795-1	Principles and framework
19795-2	Testing methodologies for technology and scenario evaluation
19795-3	Modality-specific testing
19795-4	Performance and interoperability testing of data interchange formats
19795-5	Performance of biometric access control systems
19795-6	Testing methodologies for operational evaluation

performance metrics introduced in this article are quoted from those defined in ISO/IEC 19795-1.

This article describes not only fundamental but also auxiliary performance measures of biometric systems in terms of error rates and throughput rates. These measures are mainly defined for technology evaluation, and can be easily employed for other evaluation types. Most of the measures introduced in this article are cited from ISO/IEC IS 19795-1 [1] and 19795-2 [2]. For more detailed information, readers are recommended to refer to these standards. Meanwhile, the performance measures for interoperability testing of data interchange formats and for sensor characteristics are not considered in this article.

Performance Measures

Decision errors in biometric verification or identification are due to various types of errors occurred in each process of a biometric system, sample acquisition, feature extraction, and comparison. How these fundamental errors combine to form decision errors depends upon various factors such as the number of comparisons required, either positive or negative claim of identity, and the decision policy, for example, whether the system allows multiple attempts.

Fundamental Performance Measures

The following measures are considered to be fundamental because they can be employed regardless of the types of applications of biometric systems. The failure-to-enroll and failure-to-acquire rates measure the performance of the feature extracting component, while the false match and false nonmatch rates measure that of the matching component.

- *FTE* (*failure-to-enroll rate*) is the proportion of the population for whom the system fails to complete the enrolment process. The failure-to-enroll occurs when the user cannot present the required biometric characteristic, or when the submitted biometric sample is of unacceptably bad quality. In the latter case, stricter requirements on sample quality at enrollment will increase the failure-to-enroll rate, but improve matching performance because the failure-to-enroll cases do not contribute to the failure-to-acquire rate, or matching error rates.
- *FTA* (*failure-to-acquire rate*) is the proportion of verification or identification attempts for which the system fails to capture or locate biometric samples of sufficient quality. The failure-to-acquire case occurs when the required biometric characteristic cannot be presented due to temporary illness or injury, or when either the acquired sample or the extracted features do not satisfy the quality requirements. In the latter case, stricter requirements on sample quality at acquisition will increase the failure-to-acquire rate but improve matching performance, because the failure-to-acquire cases are not included in calculating the false match and nonmatch rates.
- *FNMR* (*false nonmatch rate*) is the proportion of genuine attempt samples falsely declared not to match the template of the same characteristic from the same user submitting the sample.
- *FMR* (*false match rate*) is the proportion of zero-effort impostor attempt samples falsely declared to match the compared nonself template.

The false match and false nonmatch rates are determined by the same decision threshold value on similarity scores. By adjusting the decision threshold, there will be a trade-off between false match and false nonmatch errors. They are calculated with the number of comparisons (or attempts) and useful for evaluating the performance of a component algorithm.

Performance Measures for Verification System

Verification is one of the two major applications of biometrics, where the user makes a positive claim to an identity, features extracted from the submitted biometric sample are compared with the enrolled templates for the claimed identity, and an accept- or reject decision regarding the identity claim is returned. In evaluating the performance of biometric systems, the unit operation is a transaction, which can be a single attempt but mostly consists of multiple attempts. In this aspect, the fundamental measures, FMR and FNMR, cannot be directly applied to the overall performance evaluation of a biometric system, and the following metrics are designed for more general measures.

- *FRR* (*false reject rate*) is the proportion of verification transactions with truthful claims of identity that are incorrectly denied. When a transaction consists of a single attempt, a false rejection includes a failure-to-acquire or a false nonmatch, and the false reject rate is given by:

$$\mathrm{FRR} = \mathrm{FTA} + \mathrm{FNMR} \times (1 - \mathrm{FTA})$$

- *FAR* (*false accept rate*) is the proportion of verification transactions with zero-effort wrongful claims of identity that are incorrectly confirmed. When a transaction consists of a single attempt, a false acceptance requires a false match with no failure-to-acquire, and the false accept rate is given by:

$$\mathrm{FAR} = \mathrm{FMR} \times (1 - \mathrm{FTA})$$

A first order estimation of FRR and FAR for transactions of multiple attempts can be derived from the detection error trade-off curve. However, such estimates cannot take into account correlations in sequential attempts and in the comparisons involving the same user, and consequently can be quite inaccurate. Therefore, ISO/IEC 19795 recommends that these performance metrics shall be derived directly, using test transactions with multiple attempts as specified by the decision policy.

FRR and FAR do not include the failures occurred in enrollment. As mentioned earlier, increasing the FTE rate generally improves matching performance. For comparing the performance of biometric systems having different failure-to-enroll rates, both FRR and FAR need to be generalized so that they can take enrollment errors into account. In the following generalized FRR and FAR, a failure-to-enroll is treated as if the enrollment is completed, but all subsequent transactions by or against that enrollee fail. For a technology evaluation, the generalized FRR and FAR are defined as follows [1]:

- *GFRR* (*generalized false reject rate*) is the proportion of genuine users who cannot be enrolled, whose sample is submitted but cannot be acquired, or who are enrolled, samples acquired, but are falsely rejected.

$$\begin{aligned}\mathrm{GFRR} &= \mathrm{FTE} + (1 - \mathrm{FTE}) \times \mathrm{FRR} \\ &= \mathrm{FTE} + (1 - \mathrm{FTE}) \times \mathrm{FTA} + (1 - \mathrm{FTE}) \\ &\quad \times (1 - \mathrm{FTA}) \times \mathrm{FMR}\end{aligned}$$

- *GFAR* (*generalized false accept rate*) is the proportion of impostors who are enrolled, samples acquired, and falsely matched.

$$\begin{aligned}\mathrm{GFAR} &= (1 - \mathrm{FTE}) \times \mathrm{FAR} \\ &= (1 - \mathrm{FTE}) \times (1 - \mathrm{FTA}) \times \mathrm{FMR}\end{aligned}$$

Performance Measures for Identification System

In identification, compared with verification, the user presents a biometric sample without any claim of identity, and a candidate list of identifiers are returned as a result of matching the user's biometric features with all the enrolled templates in a database. Identification has two cases: while the closed-set identification always returns a nonempty candidate list, assuming that all the users are enrolled in the database, the open-set identification may return an empty candidate list because some potential users are not enrolled.

- *CIR* (*correct identification rate*) is the proportion of identification transactions by users enrolled in the system in which the user's correct identifier is among those returned. The identification rate at rank r is the probability that a transaction by a user enrolled in the system includes that user's true identifier within the top r matches returned. When a single point identification rank is reported, it should be referenced directly to the database size.
- *FNIR* (*false-negative identification-error rate*) is the proportion of identification transactions by users enrolled in the system in which the user's correct identifier is not among those returned.

$$\mathrm{FNIR} = \mathrm{FTA} + (1 - \mathrm{FTA}) \times \mathrm{FNMR}$$

- *FPIR* (*false-positive identification-error rate*) is the proportion of identification transactions by users not enrolled in the system, where a nonempty list of identifiers is returned. For a template database of the size N, FPIR is given as:

$$\mathrm{FPIR} = (1 - \mathrm{FTA}) \times \{1 - (1 - \mathrm{FMR})^N\}$$

Other Performance Measures

Besides the above performance measures from ISO/IEC 19795-1, the following measures have been defined

for more accurate evaluation of performance of biometric systems and employed in many biometric algorithm contests such as FVC's [8–10].

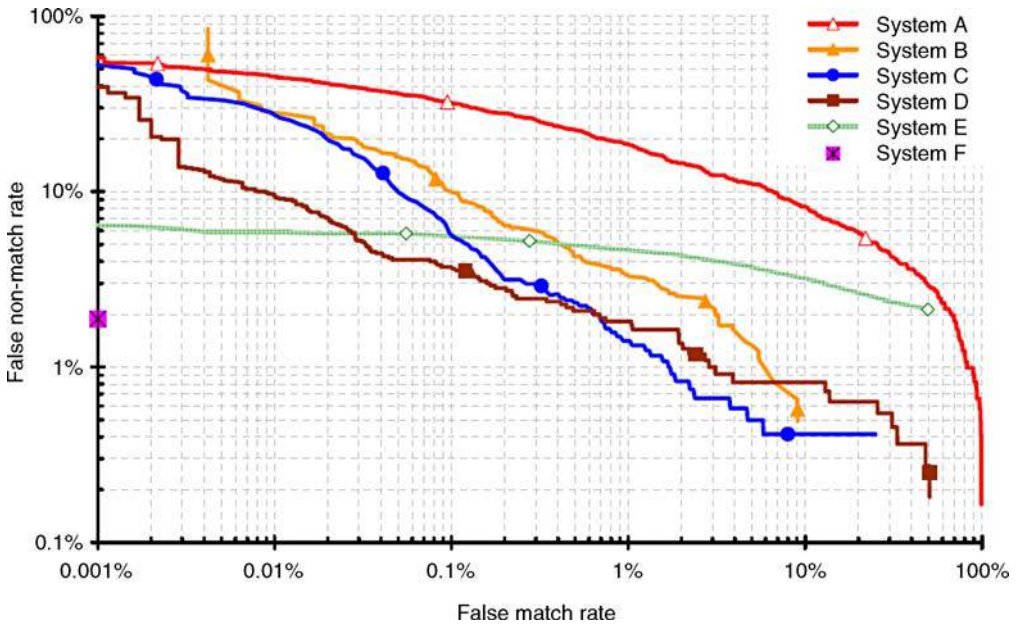
- *Genuine score distribution* and *Impostor score distribution* are computed and graphically reported to show how the algorithm "separates" the two classes.
- *EER* (*equal error rate*) is computed as the point where FNMR=FMR. In practice, the matching score distributions are not continuous and a crossover point might not exist.
- *EER** is the value that EER would take if the matching failures were excluded from the computation of FMR and FNMR.
- *FMR100* is the lowest FNMR for FMR $\leq$ 1%.
- *FMR1,000* is the lowest FNMR for FMR $\leq$ 0.1%.
- *ZeroFMR* is the lowest FNMR at which no False Matches occur.
- *ZeroFNMR* is the lowest FMR at which no False NonMatches occur.
- *Average enroll time* is the average CPU time for a single enrollment operation.
- *Average match time* is the average CPU time for a single match operation between a template and a test sample.

Graphic Performance Measures

When presenting test results, the matching or decision-making performance of biometric systems are graphically represented using Detection Error Trade-off (DET), Receiver Operating Characteristics (ROC), or Cumulative Match Characteristic (CMC) curves.

DET Curve

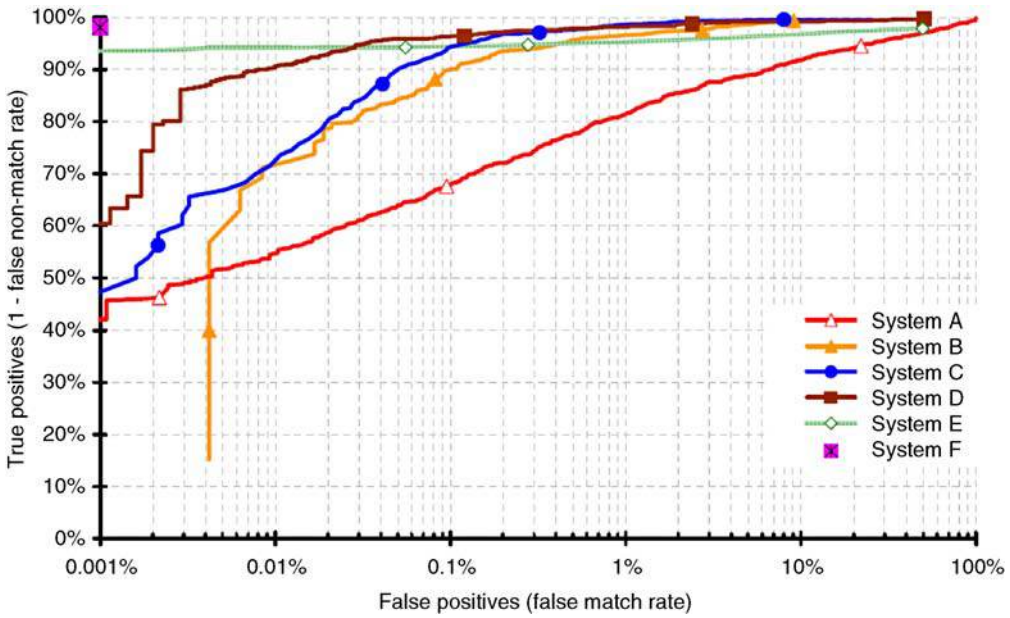
DET curves are used to plot matching error rates (FNMR against FMR), decision error rates (FRR against FAR), and open-set identification error rates (FNIR against FPIR). The DET curve is a modified ROC curve which plots error rates on both axes (false positives on the *x*-axis and false negatives on the *y*-axis). For example, in Fig. 1, each DET curve is generated by varying the value of the decision threshold. If the threshold is set to a higher value in order to decrease the false acceptances, the false rejections will increase. On the contrary, if the threshold is set to a lower value, the false rejections will decrease with the increase in false acceptance.



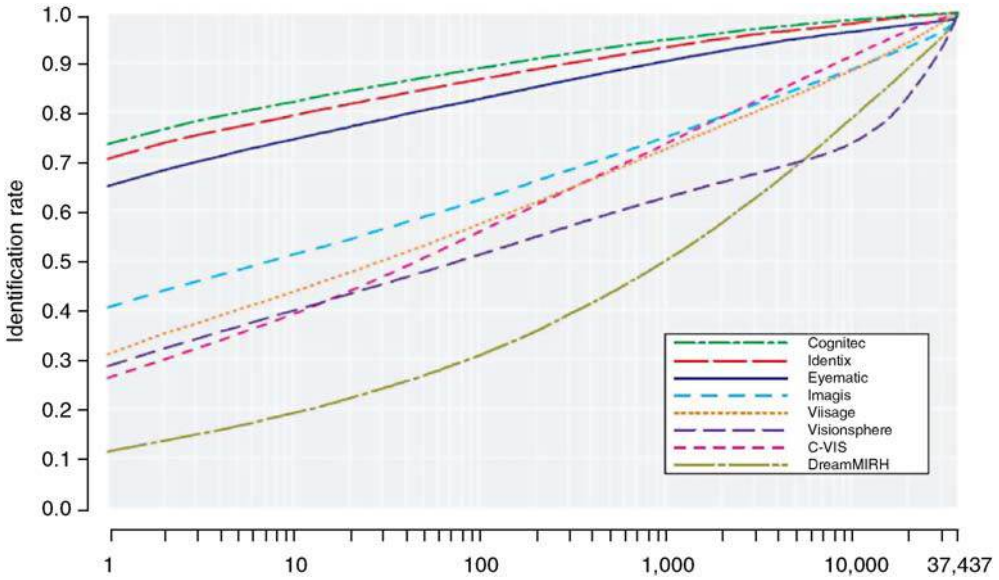
Performance Measures. Figure 1 Example set of DET curves [1].

ROC Curve

ROC curves are a traditional method for summarizing the performance of imperfect diagnostic, detection, and pattern-matching systems. ROC curves are threshold-independent, allowing performance comparison of different systems under similar conditions, or of a single system under differing conditions. ROC curves may be used to plot matching algorithm performance (1−FNMR against FMR), end-to-end verification system performance (1−FRR against FAR), as well as open-set identification system performance (CIR against FPIR). Figure 2 shows an example of ROC curves for comparing the performance of a set of fingerprint matching



Performance Measures. Figure 2 Example set of ROC curves [1].



Performance Measures. Figure 3 Example set of CMC curves [7].

algorithms. An ROC curve is a plotting of the rate of false positives (i.e., impostor attempts accepted) on the *x*-axis against the corresponding rate of true positives (i.e., genuine attempts accepted) on the *y*-axis plotted parametrically as a function of the decision threshold.

CMC Curve

For closed-set identification applications, performance results are often illustrated using a cumulative match characteristic curve. Figure 3 shows an example of CMC curves for comparing the performance of a set of face identification systems. These curves provide a graphical presentation of identification test results and plots rank values on the *x*-axis with the corresponding probability of correct identification at or below that rank on the *y*-axis.

Throughput Rates

Throughput rates represent the number of users that can be processed per unit time, based on both computational speed and human–machine interaction. These measures are generally applicable to all biometric systems and devices. Attaining adequate throughput rates is critical to the success of any biometric system. For verification systems, throughput rates are usually controlled by the speed of user interaction with the system in the process of submitting a biometric sample of good quality. For identification systems, they can be heavily impacted by the computer processing time required to compare the acquired sample with the database of enrolled templates. Hence, depending upon the type of a system, it may be appropriate to measure the interaction times of users with the system and also the processing rate of the computational hardware. Actual benchmark measurement of computer processing speed is covered elsewhere and is considered outside the scope of this article [13].

Related Entries

- ▶ CMC Curve
- ▶ DET Curve
- ▶ Biometric Sample Quality
- ▶ Influential Factors to Performance
- ▶ Interoperable Performance
- ▶ ROC Curve

References

1. ISO/IEC JTC1/SC37 IS19795-1: Biometric Performance Tasting and Reporting- Part 1: Principles and Framework (2006)
2. ISO/IEC JTC1/SC37 FDIS 19795-2: Biometric Performance Tasting and Reporting- Part 2: Testing methodologies for technology and scenario evaluation (2006)
3. Phillips, P.J., Martin, A., Wilson, C.L., Przybocki, M.: An Introduction to Evaluating Biometric Systems. IEEE Computer Magazine (2000)
4. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face-recognition algorithms. Image Vis Comput, **16**(5), 295–306 (1998)
5. Phillips, P.J., Moon, H., Rizvi, S., Rauss, P.: The FERET evaluation methodology for face recognition algorithms. IEEE Trans. PAMI, **22**, 1090–1104 (2000)
6. Blackburn, D., Bone, M., Phillips, P.J.: Face recognition vendor test 2000. Technical report. http://www.frvt.org (2001). Accessed 19 Aug, 2008
7. Phillips, P.J., Grother, P., Micheals, R., Blackburn, D., Tabassi, E., Bone, J.: Face Recognition Vendor Test 2002: Evaluation Report. Technical Report NISTIR 6965, National Institute of Standards and Technology (2003)
8. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2000: fingerprint verification competition. IEEE Trans. PAMI, **24**(3), 402–412 (2000)
9. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2002: In: Second Fingerprint Verification Competition. Proceedings of the 16th ICPR, vol. 3, pp. 811–814 (2002)
10. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2004: Third Fingerprint Verification Competition. Proceedings of the ICBA Springer LNCS, **3072**, pp. 1–7 (2004)
11. NIST Speaker Recognition Evaluations. Available online at http://www.nist.gov/speech/tests/spk/
12. Martin, A., Przybocki, M.: The NIST 1999 speaker recognition evaluation an overview. Digit. Signal Process. **10**, 1–18 (2000)
13. Mansfield, A.J., Wayman, J.L.: Best Practices in Testing and Reporting Performance of Biometric Devices, Version 2.01, NPL Report CMSC 14/02 (2002)

Performance Metrics

- ▶ Performance Measures

Performance of Biometric Quality Measures

▶ Biometric Sample Quality

Performance Testing

▶ Performance Evaluation, Overview

Performance Testing Methodology Standardization

Michael Thieme
International Biometric Group, New York, NY, USA

Synonym

Biometric performance evaluation standardization

Definition

Performance testing methodology standards define processes for test planning, hardware and software configuration and calibration, data collection and management, enrollment and comparison, performance measurement and reporting, and documenting the statistical significance of test results. The application of performance testing standards enables meaningful measurement, prediction, and comparison of biometric systems' enrollment rates, accuracy, and throughput. Interoperability of biometric data elements acquired or generated through different components can also be quantified through standardized performance tests. Standardized performance testing methodologies have been developed for technology tests, in which algorithms process archived biometric data; scenario tests, in which biometric systems collect and process data from test subjects in a specified application; and operational tests, in which a biometric system collects and processes data from actual system users in a field application.

Motivation for the Development of Biometric Performance Evaluation Standards

The development of biometric performance testing standards has been driven by the need for precise, reliable, and repeatable measurement of biometric system accuracy, capture rates, and throughput. Match rates, enrollment and acquisition rates, and throughput are central considerations for any organization deciding whether to deploy biometrics or determining which modalities and components to implement. Organizations need to know whether a claimed performance level for System A can be compared to a claimed performance level for System B; if test conditions varied between two evaluations, comparison of the same performance metrics may be useless. For example, if a vendor claims that its system delivers a false match rate (FMR) of 0.01%, for example, a potential deployer might ask:

1. How many test subjects and samples were used to generate this figure?
2. What was the composition of the test population whose data was used?
3. How much time elapsed between enrollment and verification?
4. Were all comparisons accounted for, or were some samples discarded at some point in the test?
5. What is the statistical significance of the claimed error rate?
6. What were the corresponding false non-match rate (FNMR) and failure to enroll rate (FTE) at this operating point?
7. Was the algorithm tuned to perform for a specific application or test population?
8. How were test subjects trained and guided?
9. How were errors discovered?

Organizations also need to understand biometric performance evaluation standards in order to properly specify performance requirements. A lack of understanding of biometric performance testing often leads organizations to specify requirements that cannot be validated through testing.

Once an organization has decided to deploy a biometric system, standardized performance testing methods are no less important. Organizations must properly calibrate systems prior to deployment and monitor system performance once operational. This

calibration and monitoring requires standardized approaches to data collection, data management, processing, and results generation.

Types of Biometric Performance Testing Standards

Biometric performance tests are typically categorized as technology tests, scenario tests, or operational tests. These test types share commonalities – addressed in framework performance testing standards – but also differ in important ways.

▶ Technology tests are those in which biometric algorithms enroll and compare archived (i.e., previously-collected) data. An essential characteristic of technology testing is that the test subject is not "in the loop" – the test subject provides data in advance, and biometric algorithms are implemented to process large quantities of test data. Technology tests often involve cross-comparison of hundreds of thousands of biometric samples over the course of days or weeks. Methods of executing and handling the outputs of such cross-comparisons are a major component of technology-based performance testing standards. Technology tests are suitable for evaluation of both verification- and identification-based systems, although most technology tests are verification-based. Technology testing standards accommodate evaluations based on biometric data collected in an operational system as well as evaluations based on biometric data collected for the specific purpose of testing. Technology tests based on operational data are often designed to validate or project the performance of a fielded system, whereas technology tests based on specially-collected data are typically more exploratory or experimental.

▶ Scenario tests are those in which biometric systems collect and process data from test subjects in a specified application. An essential characteristic of scenario testing is that the test subject is "in the loop," interacting with capture devices in a fashion representative of a target application. Scenario tests evaluate end-to-end systems, inclusive of capture device, quality validation software, enrollment software, and matching software. Scenario tests are based on smaller sample sizes than technology tests due to the costs of recruiting and managing interactions with test subjects (even large scenario tests rarely exceed more than several hundred test subjects). Scenario tests are also limited in that there is no practical way to standardize the time between enrollment- and recognition-phase data collection. This duration may be days or weeks, depending on the accessibility of test subjects. Scenario-based performance testing standards have defined the taxonomy for interaction between the test subject and the sensor; this taxonomy addresses presentations, attempts, and transactions, each of which describes a type of interaction between a test subject and a biometric system. This is particularly important in that scenario testing is uniquely able to quantify "level of effort" in biometric system usage; level of effort directly impacts both accuracy and capture rates.

▶ Operational tests are those in which a biometric system collects and processes data from actual system users in a field application. Operational tests differ fundamentally from technology and scenario tests in that the experimenter has limited control over data collection and processing. Because operational tests should not interfere with or alter the operational usage being evaluated, it may be difficult to establish ground truth at the subject or sample level. As a result, operational tests may or may not be able to evaluate false accept rates (FAR), false reject rates (FRR), or failure to enroll rates (FTE); instead they may only be able to evaluate acceptance rates (without distinction between genuine and impostor) and operational throughput. One of the many challenges facing developers of operational testing standards is the fact that each operational system differs in some way from all others, such that defining commonalities across all such tests is difficult to achieve. It is therefore essential that operational performance test reports specify which elements were measurable and which were not. Operational tests may also evaluate performance over time, such as with a system in operation for a number of months or years.

In a general sense, as a given biometric technology matures, it passes through the cycle of technology, scenario, and then operational testing. Biometric tests may combine aspects of technology, scenario, and operational testing. For example, a test might combine controlled, "online" data collection from test subjects (an element of scenario testing) with full, "offline" comparison of this data (an element of technology testing). This methodology was implemented in iris recognition testing sponsored by the US Department of Homeland Security in 2005 [1].

Elements Required in Biometric Performance Testing Standards

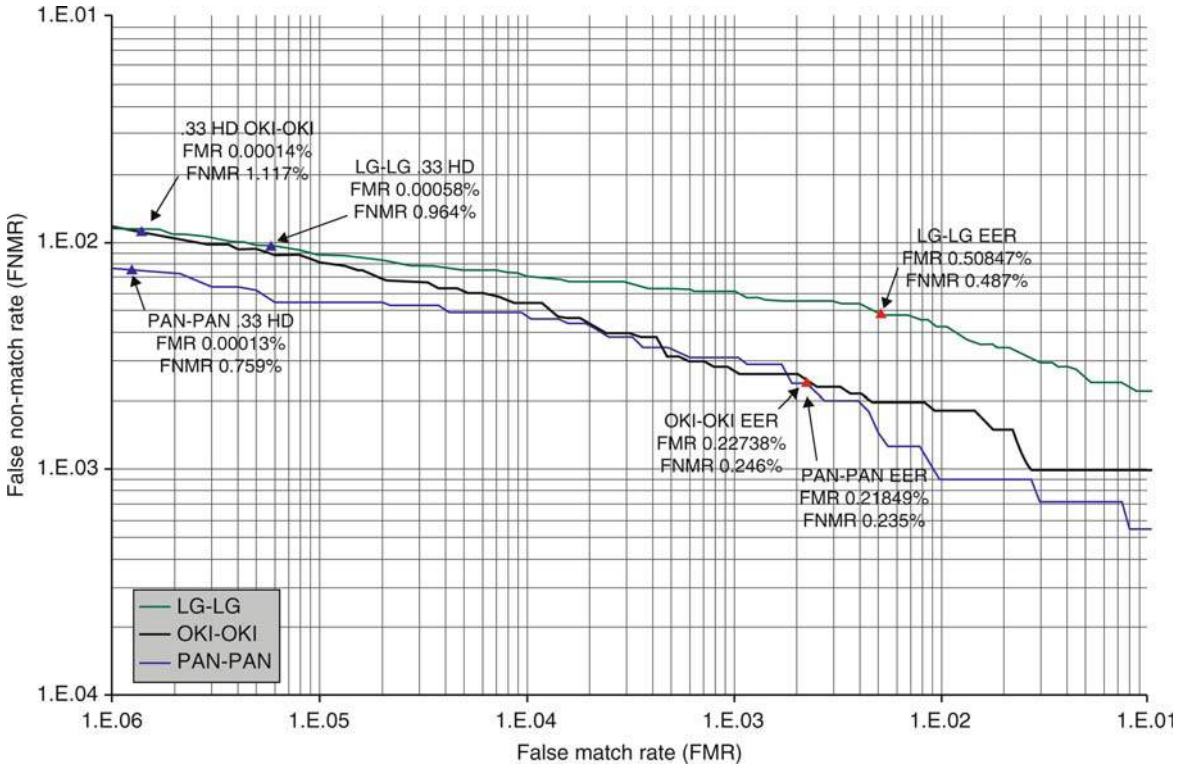
Biometric performance testing standards address the following areas:

1. *Test planning*, including requirements pertaining to test objectives, timeframes, controlling test variables, data collection methods, and data processing methods.
2. *Hardware and software configuration and calibration*, including requirements pertaining to algorithm implementation and device settings.
3. *Data collection and management*, including requirements pertaining to identification of random and systematic errors, collection of personally-identifiable-data, and establishing ground truth.
4. *Enrollment and comparison processes*, including requirements pertaining to implementation of genuine and impostor attempts and transactions for identification and verification.
5. *Calculation of performance results*, including formulae for calculating match rates, capture rates (FTA and FTE), and throughput rates.
6. *Determination of statistical significance*, including requirements pertaining to confidence interval calculation and reporting.
7. *Methodology and results reporting*, including requirements pertaining to test report contents and format.

One of the major accomplishments of biometric testing standards has been to specify the manner in which the tradeoff between FMR and FNMR is rendered in chart form.

Verification system performance can be rendered through detection error trade-off (DET) curves or receiver operating characteristic (ROC) curves. DET curve plots false positive and false negative error rates on both axes (false positives on the x-axis and false negatives on the y-axis), as shown below. ROC curves plot of the rate of false positives (i.e., impostor attempts accepted) on the x-axis against the corresponding rate of true positives (i.e., genuine attempts accepted) on the y-axis plotted parametrically as a function of the decision threshold (Fig. 1).

Identification system performance rendering is slightly more complex, and is dependent on whether the test is open-set or closed-set.



Performance Testing Methodology Standardization. **Figure 1** Detection error tradeoff (DET) curve.

Depending on the type of test (technology, scenario, operational), certain elements will be emphasized more than others, and results of the presentation will differ based on whether a test implements verification or identification.

Published Standards and Ongoing Efforts – International and National Activities

Several biometric performance testing standards have been published, and many additional biometric performance testing standards are in development. This section discusses ISO/IEC 19795-1, -2, and -3. ISO/IEC 19795-4 is discussed below under *Performance testing and interoperability*. These four standards are listed in the Registry of USG Recommended Biometric Standards Version 1.0, DRAFT for Public Comment, NSTC Subcommittee on Biometrics and Identity Management.

ISO/IEC 19795-1:2006 Information technology – Biometric performance testing and reporting – Part 1: Principles and framework [2] can be considered the starting point for biometric performance testing standardization. This document specifies how to calculate metrics such as false match rates (FMR), false non-match rates (FNMR), false accept rates (FAR), false reject rates (FRR), failure to enroll rates (FTE), failure to acquire rates (FTA), false positive identification rates (FPIR), and false negative identification rates (FNIR). 19795-1 treats both verification and identification testing, and is agnostic as to modality (e.g., fingerprint, face recognition) and test type (technology, scenario, operational).

ISO/IEC 19795-2:2007 Information technology – Biometric performance testing and reporting – Part 2: Testing methodologies for technology and scenario evaluation [3] specifies requirements for technology and scenario evaluations, described above. The large majority of biometric tests are of one of these two generic evaluation types. 19795-2 builds on 19795-1, and is concerned with "development and full description of protocols for technology and scenario evaluations" as well as "execution and reporting of biometric evaluations reflective of the parameters associated with biometric evaluation types." [4] 19795-2 specifies which performance metrics and associated data must be reported for each type of test. The standard also specifies requirements for reporting on decision policies whereby enrollment and matching errors are declared.

ISO/IEC TR 19795-3:2007 Information technology – Biometric performance testing and reporting – Part 3: Modality-specific testing [5] is a technical report on modality-specific considerations. 19795-1 and -2 are modality-agnostic (although they are heavily informed by experts' experience with fingerprint, face, and iris recognition systems). 19795-3, by contrast, reports on considerations specific to performance testing of fingerprint, face, iris, hand geometry, voice, vein recognition, signature verification, and other modalities. These considerations are important to deployers and system developers, as test processes vary from modality to modality. For example, in iris recognition testing, documenting biometric-oriented interaction between the subject and sensor is a central consideration to both usability and accuracy; in face recognition testing, capture variables are much less likely to impact performance.

Within the US, three biometric performance testing standards were developed prior to publication of the ISO IEC standards discussed above.

1. ANSI INCITS 409.1-2005 Information Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework
2. ANSI INCITS 409.2-2005 Information Technology – Biometric Performance Testing and Reporting – Part 2: Technology Testing and Reporting
3. ANSI INCITS 409.3-2005 Information Technology – Biometric Performance Testing and Reporting – Part 3: Scenario Testing and Reporting

Performance Testing and Interoperability

ISO/IEC 19795-4 Biometric performance testing and reporting – Part 4: Interoperability performance testing specifies requirements for evaluating the accuracy and interoperability of biometric data captured or processed through different vendors' systems. The standard, whose publication is anticipated in 2008, can be used to evaluate systems that collect data in accordance with 19794-N data exchange standards. 19795-4 helps

quantify the accuracy of standards' generic data representations relative to those of proprietary solutions. For example, is System A less accurate when processing standardized data than when processing proprietary data? Can System A reliably process standardized data from System B, and vice versa? 19795-4 contemplates online (scenario), offline (technology), and hybrid (scenario and technology) tests.

19795-4 is perhaps the highest-visibility performance testing standard due to the close relationship it bears with 19794-N standards. Data interchange standards (and conformance to these standards) have been the focus of much of the international community's efforts in biometric standardization. 19795-4 specified methods through which the adequacy of these standards can be implicitly or explicitly evaluated, leading to revisions or improvements in the standards where necessary.

Related End-User Testing Activities

Test efforts that have asserted compliance with published performance testing standards include but are not limited to the following:

1. NIST Minutiae Interoperability Exchange Test (MINEX) [6], asserts compliance with ISO/IEC 19795-4, Interoperability performance testing.
2. U.S. Transportation Security Administration Qualified Product List (QPL) Testing [7], asserts compliance with ANSI INCITS 409.3, Scenario Testing and Reporting.
3. NIST Iris Interoperability Exchange Test (IREX 08) [8], asserts compliance with ISO/IEC 19795-4, Interoperability performance testing.

Current and Anticipated Customer Needs in Biometric Performance Testing

One challenge facing biometric performance testing standardization is that of successfully communicating performance results to non-specialist customers (e.g., managers responsible for making decisions on system implementation). To successfully utilize even standards-compliant test reports, the reader must learn a range of acronyms, interpret specialized charts, and understand the test conditions and constraints. The "so what"? is not always evident in biometric performance test reports. This is particularly the case when trying to graphically render error bounds and similar uncertainty indicators associated with performance test results.

A difficult-to-avoid limitation of biometric performance testing standards is that tests results will differ based on test population, collection processes, data quality, and target application. In other words, a systems' error rate is not necessarily a reflection of its robustness, even if a test conforms to a standard.

Gaps in Standards Development

Performance has been defined somewhat narrowly in the biometric standards arena, most likely because the first-order consideration for biometric technologies has been the ability to reduce matching error rates. The traditional focus on matching error rates – particularly FMR – in biometric performance testing may be considered disproportionate in the overall economy of biometric system performance. As accuracy, enrollment, and throughput rates improve with the maturation of biometric technologies, development of performance testing standards may be required in areas such as usability, reliability, availability, and resistance to deliberate attacks. For example, the number of "touches" required to negatively impact match rates associated with images captured from a fingerprint sensor could be the subject of a performance testing standard.

An additional gap in biometric performance testing standards is in testing under non-mainstream conditions, such as with devices exposed to cold or to direct sunlight, or with untrained populations and/or operators. Many tests are predicated on controlled-condition data collection, though biometric applications for population control or military operations are often highly uncontrolled.

Conformance testing is a third gap in performance testing standards. The international community is working on methods for validating that test reports and methodologies conform to published standards. Certain elements can be validated in an automated fashion, such as the presence of required performance

data; other elements, such as those that describe how testing was conducted, may be reliant on test lab assertions.

Role of Industry/Academia in the Development of the Required Testing Methods

Biometric performance tests predated the development of standardized methodologies by several years. Government and academic researchers and scientists gradually refined performance testing methods in the early 1990s, with many seminal works performed in voice recognition and fingerprint. The National Biometric Test Center at San Jose State University [9] was an early focal point of test methodology development. Today, leading developers of performance testing standards include the National Institute of Standards and Technology (NIST) [10], an element of the US Department of Commerce, and the UK National Physical Laboratory (NPL) [11].

Biometric vendors bring to bear considerable expertise on performance testing, having unparalleled experience in testing their sensors and algorithms. However, vendors are not highly motivated to publish comprehensive, standards-compliant performance tests. Speaking generally, vendors are most interested in practical answer to questions such as, how many test subjects and trials are necessary to assert a FMR of 0.1%? Biometric services companies (e.g., consultancies and systems integrators) also support government agencies in standardized performance test design and execution.

Summary

Biometric performance testing standards enable repeatable evaluations of biometric algorithms and systems in controlled lab and real-world operational environments. Performance testing standards are central to successful implementation of biometric systems, as government and commercial entities must be capable of precisely measuring the accuracy and usability of implemented systems. Deployers must also be able to predict future performance as identification systems grow larger and as transaction volume increases.

References

1. http://www.biometricgroup.com/reports/public/ITIRT.html
2. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41447
3. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41448
4. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41448
5. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41449
6. http://fingerprint.nist.gov/minex/
7. http://www.tsa.gov/join/business/biometric_qualification.shtm
8. http://iris.nist.gov/irex/IREX08_conops_API_v2.pdf
9. http://www.engr.sjsu.edu/biometrics/
10. http://www.itl.nist.gov/div893/biometrics/standards.html
11. http://www.npl.co.uk

Perpetrator Identification

▶ Gait, Forensic Evidence of

Personal Data

Personal data is any information that may be used, either individually or when combined with other data, to identify a person. For example, a name and address are usually sufficient to identify a person, and so this data pair would be considered personal data. A more subtle example is hobbies and Postal Code obtained from a retail liquor store's customer database. By themselves, these data items would not be uniquely identifying, but when linked with another database, say, that of an iguana owners club, and add to that the common sense knowledge that iguana make rare pets, it becomes almost certain who patronized the liquor store. In other words, any data that could potentially reveal the identity of a person must be treated as personal data.

▶ Privacy Issues

Personal Information Search

► Background Checks

Person-Independent Model

► Universal Background Models

Phase

Every complex number $a+bi$ can be expressed in polar coordinate form as Ae^{ip}, where A is the amplitude and p is the phase of the complex number. The phase p can be computed by arctan (b/a). Phase is always in the range between zero and 2π.

► Iris Recognition, Overview

Phoneme

The phoneme is to spoken language what the letter is to written language – the representation of an individual speech sound. The English word "sick" comprises the three phonemes /s-I-k/, whereas "thick" consists of /θ-I-k/, "sack" consists of /s-æ-k/, and sit consists of /s-I-t/. The phonetic symbols, which can be used for the sounds of any language, are defined in the International Phonetic Alphabet. Strictly speaking, the phoneme is an abstract concept used in linguistics, and phonemes often correspond to complex compound sounds. For example, the phoneme /b/ (as in "big") is produced by the speaker first closing the lips and then releasing the built-up air pressure creating a plosive sound. The period of lip closure, the plosion and the transitions to and from the previous and subsequent phonemes are all identifiable as distinct events in the audio stream.

► Liveness Assurance in Voice Authentication
► Voice Sample Synthesis

Photogrammetry

Photogrammetry is an examination of two images taken from different positions to identify three-dimensional points on an object or face represented in two dimensions. It is akin to methods of stereoscopy.

► Face, Forensic Evidence of

Photography for Face Image Data

Ted Tomonaga
Konica Minolta Technology Center, Inc., Tokyo, Japan

Synonyms

Face photograph; Facial photograph; ID photograph; Photography guidelines; Photometric guidelines

Definition

Face photography is used in passports, visas, driver licenses, or other identification documents. Face photographs can be used by human viewers or by automated face recognition systems, either for confirmation of a claimed identity (usually termed verification) or, by searching a database of face images, for determining the possible identity of an individual (usually termed identification). ISO/IEC 19794-5 defines a standard data format for digital face images to allow interoperability among face recognition systems, government agencies, and other creators and users of face images. In addition to image quality factors such as ► resolution, ► contrast, and ► brightness, many other factors affect face recognition accuracy, including subject positioning, ► pose and ► expression, ► illumination

uniformity, and the use of eyeglasses or makeup, as well as the time difference between two photographs being compared.

Introduction

Following the selection of a digital face image as the primary biometric technology for use in ► ePassports by the International Civil Aviation Organization (ICAO) in 2003, machine-assisted face recognition has become widely used for various identification and verification purposes. Compared with other ► biometric technologies, face recognition is usually considered non-invasive and more socially acceptable [1]. The accuracy and speed of face recognition technology has improved considerably recently, due, in part, to a series of government-sponsored, objective competitions among companies and academic institutions producing face recognition systems [2, 3].

The face image interchange standard known as ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) 19794-5 *Biometric Data Interchange Formats – Face Image Data* was approved as an international standard by ISO/IEC JTC1 SC37 in 2005 [4]. That standard defines a data format for digital face images to allow interoperability among face-image-processing systems. However, there are many factors that affect face recognition system performance, including an individual's appearance, such as his or her facial characteristics, hair style, and accessories, and the image acquisition conditions, such as the camera's field-of-view, focus and shutter speed, depth-of-field, background, and lighting. As a consequence, many of the countries producing ePassports have their own guidelines for the production and submission of face photographs [5–8].

This chapter describes how to arrange lighting and reflective surfaces relative to the camera and subject, and provides specific advice on the acceptable amount of variation in ► illumination across the face, on how to avoid shadows on the face and background, and on the design of a user interface that can help ensure proper head positioning. Further information may be found in ISO/IEC 19794-5 Amendment 1:2007, *Conditions for taking photographs for face image data.*[9]. The amendment provides explicit guidance for the design of photographic studios, photo booths, and other sites producing conventional printed photographs or digital images of faces that may be used in passports, visas, driver licenses, or other identification documents.

Enrollment Guidelines

The use of automated face recognition requires that an input face image first be enrolled by the system, that is, the specific set of face features to be used by the recognition system must be measured and stored. Correct enrollment requires that the input image be of high quality and meet the following criteria. Note that the numbers displayed in parentheses in the following sections are the relevant subclause numbers of the ISO/IEC 19794-5 face image interchange standard.

Subject Guidelines

For purposes of enrollment in an automated face recognition system, the following general subject guidelines should be observed:

Pose angle – The subject must be in a frontal pose. Subjects should have their shoulders square towards the camera and be looking directly at the camera. There should be no rotation of the head left or right or up or down, nor should it be tilted towards either shoulder. For rotation of the head left/right and up/down (yaw and pitch) – the compliance requirement is $<\pm$ 5° (7.2.2). For head tilt (roll) – the compliance requirement is $<\pm$ 8° (7.2.2). The requirement for roll is less restrictive, because an in-plane rotation of the head can be corrected by automated face recognition systems more easily.

Face size/position adjustment – The adjustment of face size can be made, if needed, by changing the distance between the subject and the camera or by optical zoom magnification.

Neutral expression – The subject's face should be relaxed and without expression; in particular, the subject should not be smiling. That is, his/her expression should be neutral with eyes open and mouth closed.

Eyes closed/obstructed – There should be no obstruction of the eyes due to eyeglass rims, tint, or glare, bangs, eye patches, head clothing, or closed eyes (7.2.3, 7.2.11) Hats, scarves, or any other apparel that may obstruct the face should be removed.

Background – The background should be unpatterned and plain, such as a solid-color wall or cloth. The background color may be a light gray, light blue, white or off-white. The background should be separately illuminated such that there are no shadows visible on the background behind the subject's face (A.2.4.3).

Camera Guidelines

To ensure correct camera focus, color, and exposure, and minimal motion blur and geometric distortion, the following general guidelines should be observed:

Shutter speed – The shutter speed should be high enough to prevent motion blur (1/60–1/250 s), unless electronic flash is the predominant source of illumination.

Color balance – The image color balance should reflect natural colors with respect to expected skin tones. This value can be affected by inappropriate white balancing or red-eye (7.3.4).

Brightness exposure/contrast – Exposure should be checked with an exposure meter. Gradations in skin texture should be visible, with no saturation on the face (7.3.2).

Camera-to-subject distance – To ensure minimal geometric distortion, the camera-to-subject distance should be within the range of 1.2–2.5 m in a typical photo studio.

Camera–subject height – The camera should be tripod-mounted for stability. The optimum height of the camera is at the subject's eye-level. Height adjustment can be done either by using a height-adjusting stool for the subject or by adjusting the tripod's height.

Centering – Keeping the subject's face at the center of the frame is recommended. The horizontal center of face shall be between 45% and 55% of the image width (8.3.2). The vertical center of the face shall be between 30% and 50% of the image height, as measured from the top of the image (8.3.3).

Head size relative to the image size – Head width to image width ratio should be between 5:7 and 1:2 (8.3.4, 8.3.5, A.3.2.2).

Focal length of camera lens (35 mm format equivalent) – Use a normal to medium telephoto lens (50–130 mm) (under shooting distance of 1.2–2.5 m).

Resolution – The spatial resolution should be greater than about 2 pixels per mm. Resolution can easily be checked by test shooting a ruler (7.3.3). For optimum performance of a face recognition system, the number of pixels between the eyes shall be at least 90 (8.4.1) and preferably 120 (A.3.1.1).

Dynamic Range in Face – There should be at least 7 bits of intensity variation (i.e., at least 128 unique values) in the facial region after conversion to grayscale (7.4.2).

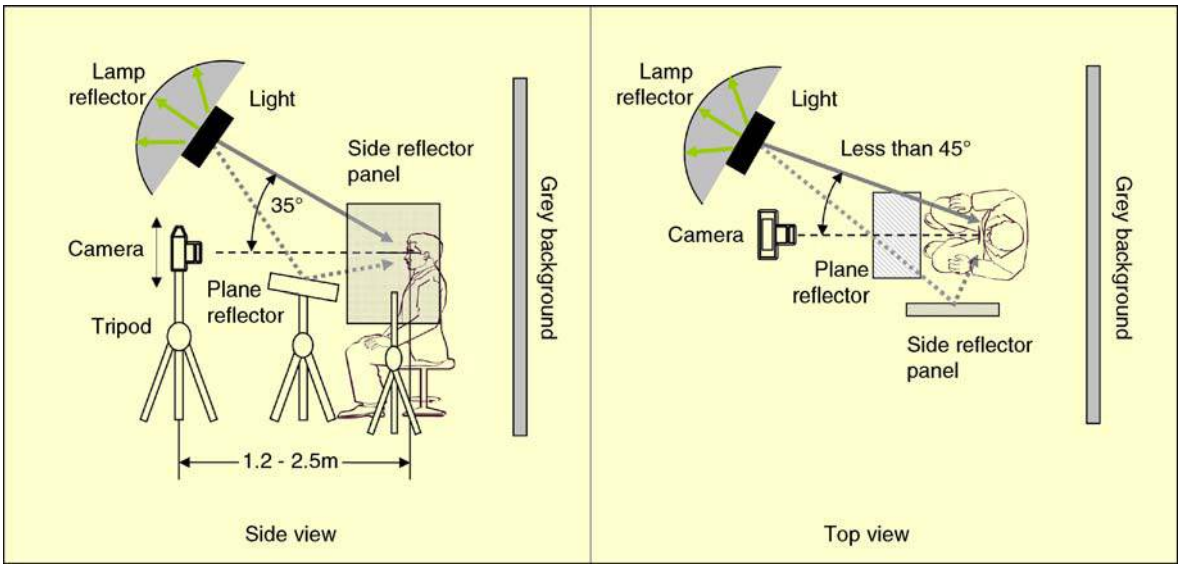
Lighting Guidelines

Artistic portraits are often taken with intentionally uneven illumination, while face photographs taken for the purposes of identification by humans or machines should display even illumination of the face. Various technical publications have reported that the use of lighting arrangements that evenly illuminate the face without producing shadows around the nose or eyes improves the accuracy of automated face recognition systems [10, 11].

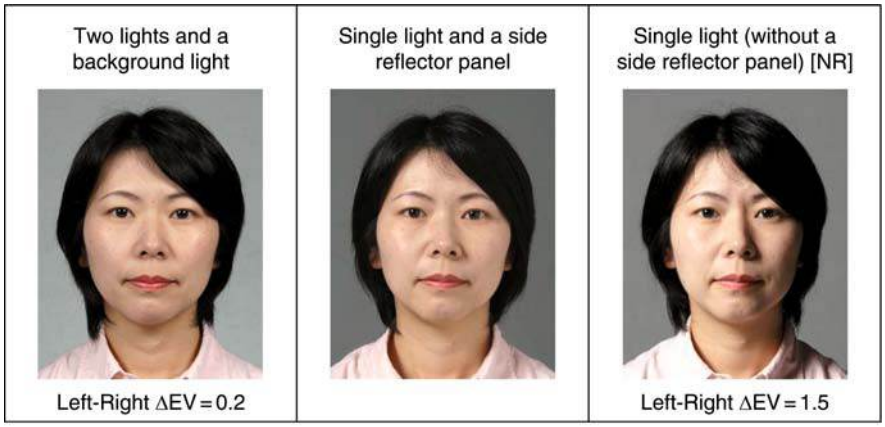
Example Configurations for a Photo Studio or Store

Typically, a photo studio or a photo store is a professionally operated facility, equipped with a film or digital camera, multiple adjustable light sources, a suitable background or backdrop cloth, and subject positioning apparatus designed to obtain high quality portraits. Some of the design considerations for a photo studio or store are described in the following paragraphs:

Lighting uniformity (No shadows and glare) on the face – A simple arrangement is a single light source and multiple reflector panels to illuminate the subject's face uniformly. The light, shown with a lamp reflector, should be placed approximately 35° above the line between the camera and the subject, and be directed towards the subject's face at a horizontal angle of less than 45° from the line (Fig. 1). Ideally there would be two diffused light sources in front of the face at 45° on either side of the camera. The maximum difference of four exposure values on the left and right sides of a face, chin, and forehead should be less than 1 EV (Fig. 2). The measurements may be made by placing an incident light meter at those four positions of a subject's face and pointing the meter towards the camera. If the values are not within 1 EV, the lights should be repositioned more symmetrically about the subject-to-camera line or additional reflective surfaces may be used to redistribute the light.



Photography for Face Image Data. **Figure 1** Single lamp arrangement for a photo studio.



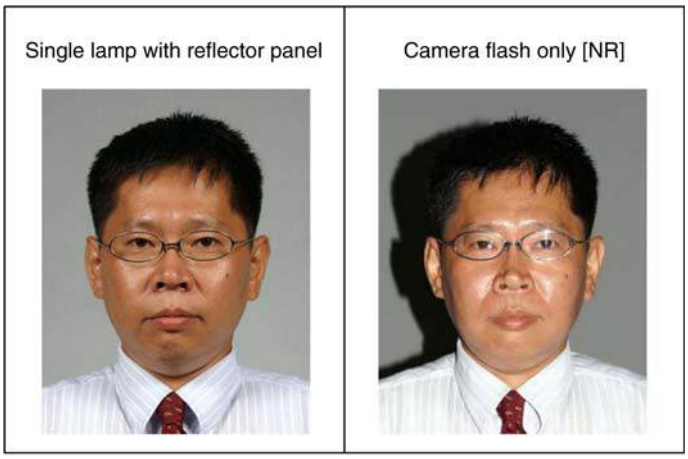
Photography for Face Image Data. **Figure 2** Evenness of illumination.

No on-camera flash – An on-camera flash should not be used. A single bare "point" light source such as a camera-mounted flash often produces "hot spots" (very bright areas on the face) and is not acceptable for imaging. (Fig. 3) The use of on-camera flash also can produce "red-eye," particularly for dark-adapted subjects.

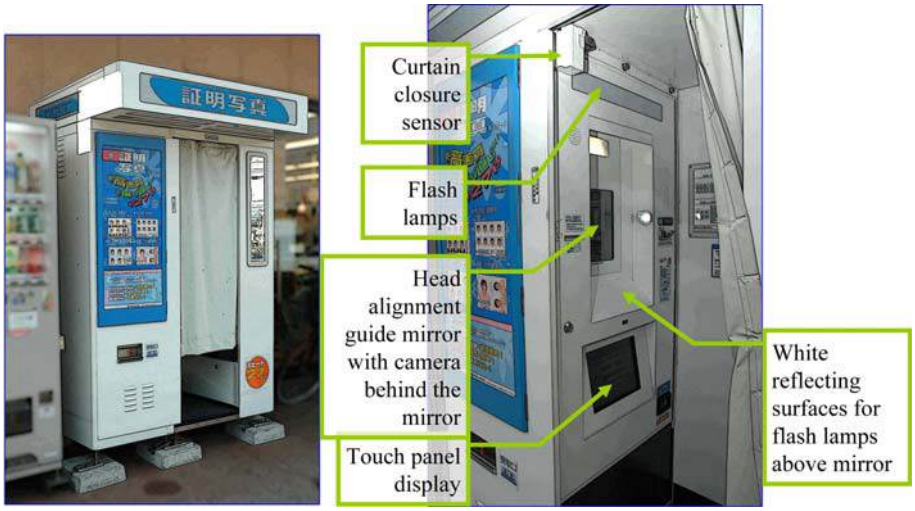
Example Configuration for a Photo Booth

A photo booth is typically a coin-operated, self-portrait photography unit, mostly used for taking ID pictures. Similar environments are used in registration offices for drivers' licenses, etc. (Figs. 4 and 5) Fig. 6 shows an example of an arrangement of lighting and a camera for a photo booth. Some of the system considerations for a photo booth are described in the following paragraphs:

Adjustment of head size, expression, etc. by monitor-GUI – There are many kinds of user interface displays for adjustment of head size and position. Figure 7 shows one of the examples of a user interface: a head positioning frame. Even with the use of a head-positioning display, an image preview should be



Photography for Face Image Data. **Figure 3** Effect of using on-camera flash note glare glasses & specular reflections from face, deep shadow on background –blending with dark hair.



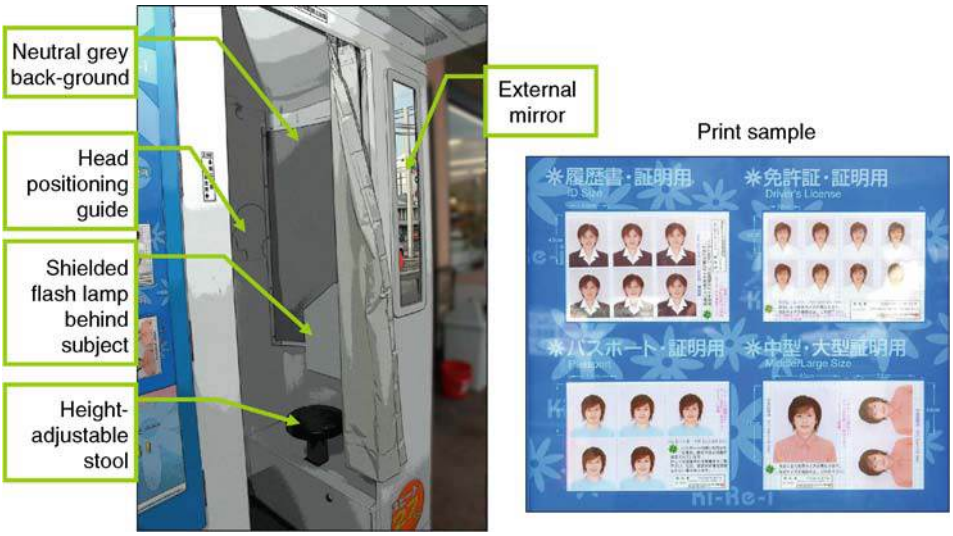
Photography for Face Image Data. **Figure 4** Example of a photo booth(1).

provided to allow thesubject to recapture the image before it's printed or written to a storage medium, in case the subject might deem his/her pose or expression unacceptable. Illustrations of acceptable poses and expressions should be provided inside the booth.

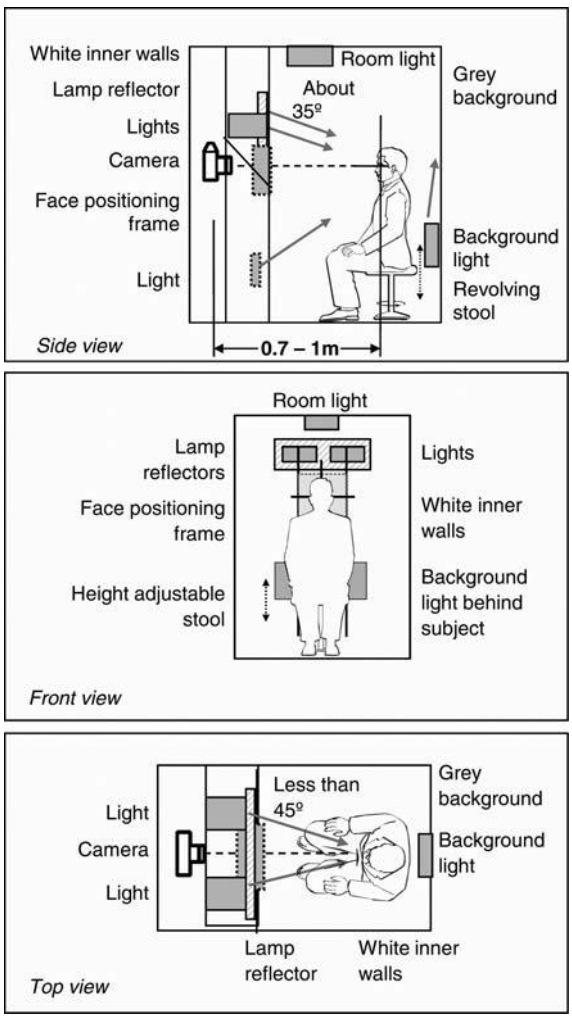
Face image quality assessment software – As an alternative to a head-positioning display, ► face detection software or ► quality assessment software that automatically sizes and centers the head within the field-of-view can be used to ensure proper head positioning. However, given that such software sometimes does not find the face position correctly, a preview image should be provided with provision for manual override of the automatically determined position.

Summary

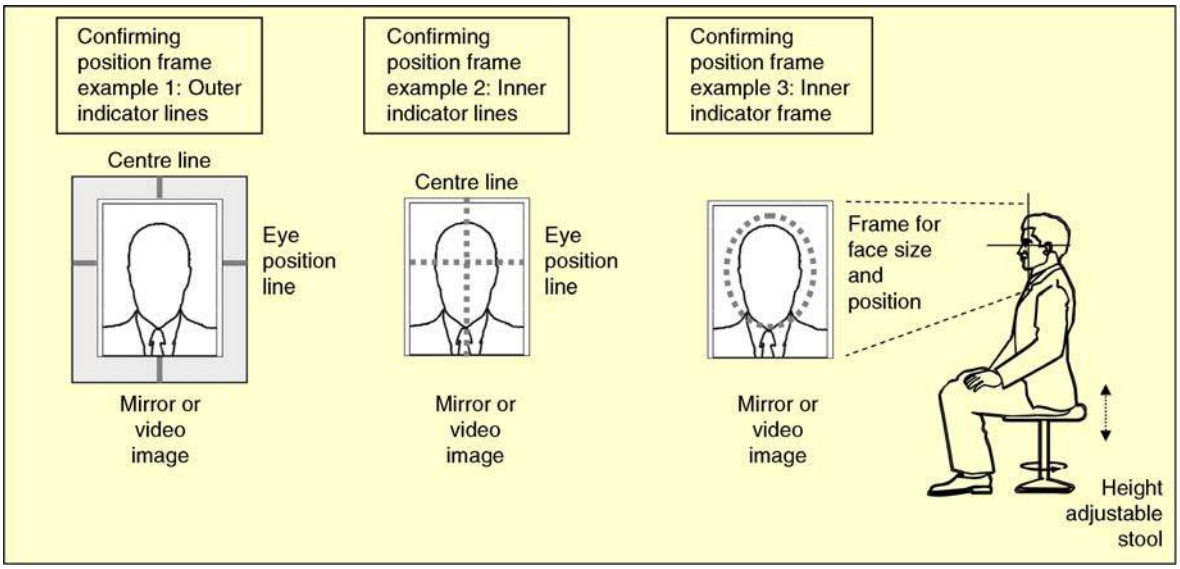
The use of proper illumination is important to obtain face photographs suitable for identification purposes. A key factor is the evenness of the illumination. Proper lighting arrangements to reduce excessive shading and



Photography for Face Image Data. **Figure 5** Example of a photo booth(2).



Photography for Face Image Data. **Figure 6** Light and camera arrangement for a photo booth.



Photography for Face Image Data. **Figure 7** User interface: Head positioning frame.

specular reflections on the face, and deep shadows on background (which can blend with thesubject's dark hair, etc.), improve the accuracy of automated face recognition systems. Other factors that can affect face image quality include pose, expression, and positioning. Increasingly, face image quality assessment software is being used to ensure high quality face images, for the preparation of not only e- , but also other ID documents.

Related Entries

- ▶ Enrollment
- ▶ Facial authentication
- ▶ Facial identification
- ▶ Facial verification

References

1. "Machine Readable Travel Documents, Part 1, Machine Readable Passports," ICAO Doc9303 6th edn. 3rd draft, 18, Nov. (2004)
2. FRVT2002. http://www.frvt.org/FRVT2002/default.htm
3. Phillips, P.J., et al.: FRVT 2006 and ICE 2006 Large-Scale Results, NISTIR 7408, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. http://face.nist.gov/frvt/frvt2006/frvt2006.htm 9. Guidelines of USA, UK, Australia and Canada
4. ISO/IEC 19794-5:2005, Information Technology – Biometric Data Interchange Formats – Part 5: Face Image Data (2005)
5. http://travel.state.gov/pdf/Photo%20Guide%2010-01-04.pdf
6. http://www.ips.gov.uk/passport/downloads/Photographers-guide-to-passport-photos-edited.pdf
7. https://www.passports.gov.au/Web/Requirements/Photos.aspx
8. http://www.ppt.gc.ca/cdn/photos.aspx; http://www.ppt.gc.ca/form/pdfs/pptc320_eng.pdf
9. ISO/IEC 19794-5:2005, Information technology – Biometric data interchange formats – Part 5: Face image data – AMENDMENT 1: Conditions for taking photographs for face image data (2007)
10. Commissioned Report: Studies on Biometric Authentication Using IC Passports. http://www.nmda.or.jp/nmda/bio/ic-passports/ic-passports.pdf
11. Junichi Sakaki, Shizuo Sakamoto, Koji Matsunaka et al.: [Special Talk] IC Passports and Systematic Evaluation toward Employing Face Authentication Technologies, IEICE (The Institute of Electronics, Information and Communication Engineers) technical report PRMU2005-92 (2005–10), pp. 51–56, (2005)
12. D'Amato, D.P.: Best Practices for Taking Face Photographs and Face Image Quality Metrics NIST Biometric Quality Workshop. http://www.itl.nist.gov/iad/894.03/quality/workshop/proc/d'amato_nist_image_quality_workshop_slides-final.pdf Accessed 8 March (2006)

Photography Guidelines

- ▶ Photography for Face Image Data

Photometric Guidelines

► Photography for Face Image Data

Photon

In classical electromagnetic theory, light and other forms of electromagnetic radiation propagate as waves. Quantum mechanics teaches us that nature is more complicated; light propagation can be modeled as either a wave or a particle, depending on the viewpoint and measurement process of the observer. A photon is a "particle" of light, in the same sense that an electron is a "particle" of electricity. When light is absorbed by a piece of silicon in an camera sensor, the model for the interaction is that photons of light are individually absorbed by the silicon and excite electron-hole pairs in the silicon that result in an electrical signal that can be processed to generate images.

► Iris Device

Physical Analogies for Ear Recognition

David J. Hurley, Mark S. Nixon
School of Electronics and Computer Science, University of Southampton, Southampton, UK

Synonyms

Convergence feature extraction: convergence

Definition

In the context of ear biometrics, Hurley et al. [1–3] have developed a pair of invertible linear transforms called the ► force field transform and the ► potential energy transform, which transform an image into a force field by pretending that pixels have a mutual attraction directly proportional to their intensities and inversely proportional to the square of the distance between them, rather like Newton's Law of Universal Gravitation. Underlying this force field, there is an associated potential energy field, which in the case of an ear, takes the form of a smooth surface with a number of peaks joined by ridges. The peaks correspond to the potential energy wells, and to extend the analogy the ridges correspond to the potential energy channels. Since the transform also happens to be invertible, all the original information is preserved, and since the otherwise smooth surface is modulated by these peaks and ridges it is argued that much of the information is transferred to these features and that therefore they should make good features for recognition. An analysis of the mechanism of this algorithmic ► field line feature extraction approach leads to a more powerful method called ► convergence feature extraction, based on the divergence of force direction revealing even more features in the form of antiwells and antichannels.

Introduction

The last 10 years has seen increasing interest in the ear as a biometric with significant contributions from computer vision researchers [1–8]. In this context the force field transform has been developed that effectively filters an ear image by convolving it with a huge inverse square kernel more than four times the size of the image, the force then being the gradient of the resulting massively smoothed image. Force field feature extraction subsequently exploits the directional properties of the force field to automatically locate ear features in the form of potential channels and wells. The force field paradigm allows us to draw upon a wealth of proven techniques from vector field calculus; for example, the divergence operator is applied to the force field direction, yielding a nonlinear operator called convergence of force direction leading to the even more powerful convergence feature extraction. The extreme kernel size results in the smoothened image having a general dome shape, which gives rise to brightness sensitivity issues. However, it is argued by showing that the field line features are hardly distorted, that this will have little overall effect, and this conclusion is borne out by including brightness variation in the

recognition tests. On the other hand, the dome shape leads to an automatic extraction advantage, and this is demonstrated by deliberately using poorly registered and poorly extracted images in recognition tests and then comparing the results with those for principal components analysis (PCA) under the same conditions, where we see that the ear images have to be accurately extracted and registered for PCA to achieve comparable results. The technique is validated by achieving a recognition rate of 99.2% on a set of 252 ear images taken from the XM2VTS face database [9]. Not only is the inherent automatic extraction advantage demonstrated but it is also shown that it performs even more favorably against PCA under variable brightness conditions, and also demonstrates its excellent noise performance by showing that noise has little effect on recognition results. Thus the technique has been validated by achieving good ear recognition results, and in the process a contribution has been made to the mounting evidence that the human ear has considerable biometric value.

Ear Feature Extraction

Force Field Feature Extraction

Here, the force field transform and algorithmic field line feature extraction are described before introducing convergence feature extraction. The mathematical concepts used can be found in basic works on electromagnetics [10] and a more detailed description of the transform can be found in [3]. Faster computation using convolution and the Fast Fourier Transform (FFT) and the question of brightness sensitivity, both theoretically and by demonstration, are considered.

The image is first transformed into a force field by treating the pixels as an array of mutually attracting particles that attract each other according to the product of their intensities and inversely to the square of the distances between them. Each pixel is assumed to generate a spherically symmetrical force field so that the total force $\mathbf{F}(\mathbf{r}_j)$ exerted on a pixel of unit intensity at the pixel location with position vector $\mathbf{r}_j$ by remote pixels with position vector $\mathbf{r}_i$ and pixel intensities $P(\mathbf{r}_i)$ is given by the vector summation,

$$\mathbf{F}(\mathbf{r}_j) = \sum_i \left\{ \begin{matrix} P(\mathbf{r}_i)\frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \forall i \neq j \\ 0 \forall i = j \end{matrix} \right\}. \quad (1)$$

The underlying energy field $E(\mathbf{r}_j)$ is similarly described by,

$$E(\mathbf{r}_j) = \sum_i \left\{ \begin{matrix} \frac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \forall i \neq j \\ 0 \forall i = j \end{matrix} \right\}. \quad (2)$$

In order to calculate the force and energy fields for the entire image, the calculations should be performed for every pixel. However, this requires the number of applications of equations 1 and 2 to be proportional to the square of the number of pixels. Therefore, for faster calculation, the process is treated as a convolution of the image with the force field corresponding to a unit value test pixel, and then invoking the Convolution Theorem to perform the calculation as a frequency domain multiplication. The result of this is then transformed back into the spatial domain. The force field equation for an n-pixel image becomes,

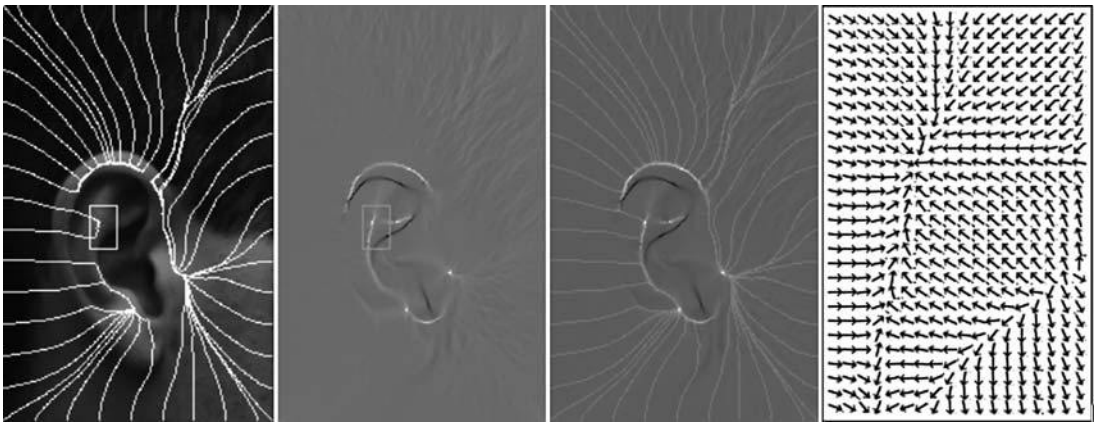
$$\text{force field} = \sqrt{n} \times \Im^{-1}[\Im(\text{unit force field}) \times \Im(\text{image})], \quad (3)$$

where $\Im$ stands for the Fourier Transform and $\Im^{-1}$ for its inverse. Listing 1

```
ff(pic) := | sr ← 2·(rows (pic) – 1), sc ← 2·(cols (pic) – 1)
           | r ← rows (pic) – 1, c ← cols (pic) – 1
           |  for rr ∈ 0.. sr
           |    for cc ∈ 0.. sc
           |      usr_{rr, cc} ← (r + c·j) – (rr + cc·j) / (|r + c·j – (rr + cc·j)|)^3
           | usr 3·rows(pic) – 3, 3·cols(pic)– 3 ← 0
           | pic 3·rows(pic) – 3, 3·cols(pic)– 3 ← 0
           | oup ← sqrt(rows (pic)·cols (pic))·icfft(cfft (usr)·cfft (pic))
           | ff← submatrix (oup, r, 2.r, c, 2·c)
```

shows how to implement this in Mathcad in which 1**j** denotes the complex operator and **cfft** and **icfft** denote the Fourier and inverse Fourier transforms, respectively. Moreover, because the technique is based on a natural force field there is the prospect of a hardware implementation in silicon by mapping the image pixels to electric charges, which would lead to very fast real time force field calculation.

Figure 1a demonstrates field line feature extraction for an ear image where a set of 44 test pixels is arranged around the perimeter of the image and allowed to



Physical Analogies for Ear Recognition. **Figure 1** Convergence field: (**a**) field lines, (**b**) convergence field, (**c**) superimposition, and (**d**) force direction.

follow the field direction so that their trajectories form field lines which capture the general flow of the force field. The test pixel positions are advanced in increments of one pixel width and the test pixel locations are maintained as real numbers, producing a smoother trajectory than if they were constrained to occupy exact pixel grid locations. Note the two obvious potential wells in the lower part of the field.

Convergence Feature Extraction

This analytical method came about as a result of analyzing in detail the mechanism of field line feature extraction. As shown in Fig. 1d, when the arrows normally used to depict a force field are replaced with unit magnitude arrows, thus modeling the directional behavior of exploratory test pixels, it becomes apparent that channels and wells arise as a result of patterns of arrows converging toward each other, at the interfaces between regions of almost uniform force direction. As this brings to mind the divergence operator of vector calculus, it was natural to investigate the nature of any relationship that might exist between channels and wells and this operator. This resulted not only in the discovery of a close correspondence between the two, but also revealed extra information corresponding to the interfaces between diverging arrows, leading to a more general description of channels and wells in the form of a mathematical function in which wells and channels are revealed to be peaks and ridges, respectively, in the function value. The new function maps the force field to a scalar field, taking the force as input and returning the additive inverse of the divergence of the force direction. The function will be referred to as the ▶ force direction convergence field $C(\mathbf{r})$ or just ▶ convergence for brevity. A more formal definition is given

$$c(\mathbf{r}) = -\mathrm{div}\,\mathbf{f}(\mathbf{r}) = -\lim_{\Delta A \to 0} \frac{\oint \mathbf{f}(\mathbf{r}) \cdot d\mathbf{l}}{\Delta A} = -\nabla.\mathbf{f}(\mathbf{r}) = -\left(\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y}\right), \quad (4)$$

where $\mathbf{f}(\mathbf{r}) = \mathbf{F}(\mathbf{r})/|\mathbf{F}(\mathbf{r})|$, ΔA is incremental area, and $d\mathbf{l}$ is its boundary outward normal. This function is real valued and takes negative values as well as positive ones where negative values correspond to force direction divergence. Listing 2

```
C(FF) := | DF ← FF / |FF|
         | for r ∈ 1... rows(DF) − 1
         |   for c ∈ 1... cols(DF) − 1
         |     dr ← Re(DF_{r,c}) − Re(DF_{r−1,c})
         |     dc ← Im(DF_{r,c}) − Im(DF_{r,c−1})
         |     C_{r,c} ← dr + dc
         | −C
```

shows a particular implementation of convergence in Mathcad where FF represents the force field and DF is the direction field.

It must also be stressed that convergence is nonlinear because it is based on force direction rather than force. This nonlinearity means that we are obliged to

perform the operations in the order shown; we cannot take the divergence of the force and then divide by the force magnitude. Div(grad/|grad|) ≠ (div grad)/|grad|. This is quite easily illustrated by a simple example using the scalar field e^x in equation 5,

$$\left\{\begin{array}{c}\text{div}(\text{grad}/|\text{grad}|)\\ \nabla\cdot\left(\frac{\nabla e^x}{|\nabla e^x|}\right)=\nabla\cdot\frac{e^x\mathbf{i}}{e^x}=\nabla\cdot\mathbf{i}=0\end{array}\right\} \neq \left\{\begin{array}{c}(\text{div grad})/|\text{grad}|\\ \frac{\nabla\cdot\nabla e^x}{|\nabla e^x|}=\frac{e^x}{e^x}=1\end{array}\right\}, \quad (5)$$

where **i** is a unit vector in the x direction. This illustrates that even though convergence looks very much like a Laplacian operator, it definitely is not.

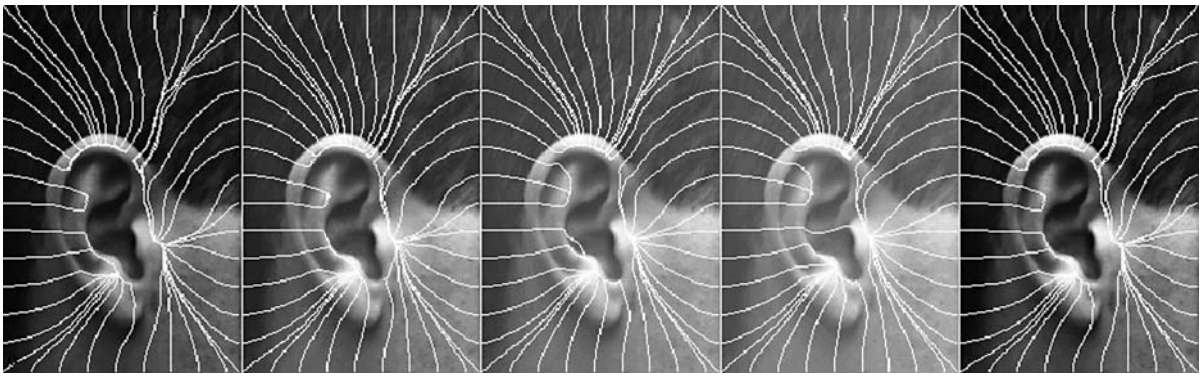
Figure 1 shows the relationship between field lines (a) and convergence (b) by merging the two fields in (c). A small rectangular section of the force direction field indicated by a small rectangular insert in Fig. 1a and b is shown magnified in Fig. 1(d). We can see clearly that channels coincide with white convergence ridges and also that wells coincide with convergence peaks, which appear as bright spots. Note the extra information in the center of the convergence map that is not in the field line map. Negative convergence values representing antichannels appear as dark bands, and positive values corresponding to channels appear as white bands. We see that the antichannels are dominated by the channels, and that the antichannels tend to lie within the confines of the channels. Note also the correspondence between converging arrows and white ridges, and between diverging arrows and black ridges. The features detected tend to form in the center of the field due to its overall dome shape, with channels and wells tending to follow intensity ridges and peaks while antichannels and antiwells tend to follow intensity troughs and hollows.

Brightness Change Analysis

Before proceeding to the next section on ear recognition, the effect of brightness change will first be analyzed by considering its effect on the energy field and then confirmed by visual experiment. Should the individual pixel intensity be scaled by a factor a and also have an additive intensity component b, we would have

$$E(\mathbf{r}_j)=\sum_i\left\{\begin{array}{c}\frac{aP(\mathbf{r}_i)+b}{|\mathbf{r}_i-\mathbf{r}_j|}\forall i\neq j\\ 0\forall i=j\end{array}\right\} = a\sum_i\left\{\begin{array}{c}\frac{P(\mathbf{r}_i)}{|\mathbf{r}_i-\mathbf{r}_j|}\forall i\neq j\\ 0\forall i=j\end{array}\right\}+\sum_i\left\{\begin{array}{c}\frac{b}{|\mathbf{r}_i-\mathbf{r}_j|}\forall i\neq j\\ 0\forall i=j\end{array}\right\}. \quad (6)$$

We see that scaling the pixel intensity by the factor a merely scales the energy intensity by the same factor a, whereas adding an offset b is more troublesome, effectively adding a pure energy component corresponding to an image with constant pixel intensity b. The effect of the offset and scaling is shown in Fig. 2 with the channels superimposed. We see that scaling by a factor of 10 in Fig. 1e has no effect as expected. The original image in Fig. 1a has a mean value of 77 and a standard deviation of 47. Figures 1b–d show the effect of progressively adding offsets of one standard deviation. At one standard deviation the effect is hardly noticeable and even at three standard deviations the change is by no means catastrophic as the channel structure alters little. We therefore conclude that operational lighting



Physical Analogies for Ear Recognition. Figure 2 Effect of additive and multiplicative brightness changes: (**a**) original, (**b**) 1 std. dev., (**c**) 2 std. devs., (**d**) 3 std. devs., (**e**) scaled ×10.

Physical Analogies for Ear Recognition. **Table 1** Comparison of force field (FFE) and PCA recognition results

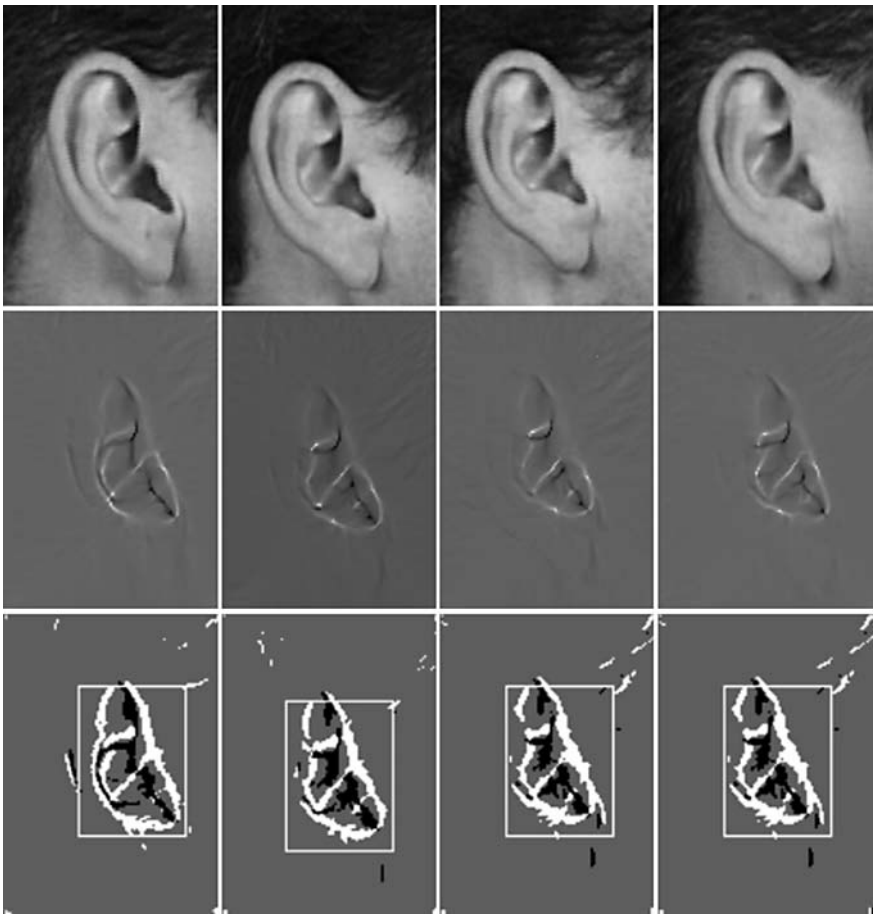
Image type	Method	Passes	Noise $20\log_{10}S/N$	Correct classification rate (%)	Brightness change by addition (std dev.)	Decidability
141×101 with deliberately poor extraction and registration	FFE	250/252	Nil	99.2	0	3.432
	FFE	251/252	18 dB	99.6	0	3.488
	FFE	249/252	12 dB	98.8	0	3.089
	FFE	241/252	6 dB	95.6	0	1.886
	FFE	250/252	Nil	99.2	1	3.384
	FFE	247/252	Nil	98.0	2	3.137
	FFE	245/252	Nil	97.2	3	2.846
	PCA	118/189	Nil	62.4	0	1.945
111 × 73 with accurate extraction and registration	PCA	186/189	Nil	98.4	0	3.774
	PCA	186/189	18 dB	98.4	0	3.743
	PCA	186/189	12 dB	98.4	0	3.685
	PCA	177/189	6dB	93.6	0	3.606
	PCA	130/189	Nil	68.8	1	1.694
	PCA	120/189	Nil	63.6	2	0.878
	PCA	118/189	Nil	62.4	3	0.476
	PCA	181/189	Nil	95.6	1 normalized	3.171
	PCA	172/189	Nil	91.0	2 normalized	1.91
	PCA	166/189	Nil	82.5	3 normalized	1.14

variation in a controlled biometrics environment will have little effect. These conclusions are borne out by the results of the corresponding recognition experiments in Table 1.

Ear Recognition

The technique was validated on a set of 252 ear images taken from 63 subjects selected from the XM2VTS face database [9] by multiplicative template matching of ternary threshold convergence maps where levels less than −1 standard deviation are mapped to −1, while those greater than one standard deviation map to +1, and those remaining map to 0. A threshold level of one standard deviation was chosen experimentally resulting in the template channel thickness as shown in Fig. 3c. This figure also shows a rectangular exclusion zone centered on the convergence magnitude centroid; the centroid of the convergence tends to be stable with respect to the ear features and this approach has the added advantage of removing unwanted outliers such as bright spots caused by spectacles. The size of the rectangle was chosen as 71 × 51 pixels by adjusting its proportions to give a good fit for the majority of the convergence maps. Note how for image 000–2, which is slightly lower than the other three, the centroid-centered rectangle has correctly tracked the template downward.

The inherent automatic extraction advantage was demonstrated by deliberately extracting or registering the ears inaccurately, in the sense that the database consists of 141 × 101 pixel images, where the ears have only an average size of 111 × 73 and are only roughly located by the eye in the center of these images. This can be seen clearly in Fig. 3a where we see a marked variation both in vertical and horizontal ear-location, and also that there is a generous margin surrounding the ears. The force field technique gives a correct classification rate (CCR) of 99.2% on this set, whereas running PCA [11] on the same set gives a result of only 62.4% but when the ears are accurately extracted by cropping to the average ear size of 111 × 73, running PCA then gives a result of 98.4%, thus demonstrating the inherent extraction advantage. The first image of the four samples from each of the 63 subjects was used in forming the PCA covariance matrix. Figure 4 shows the first four eigenvectors for the 111 × 73-pixel



Physical Analogies for Ear Recognition. **Figure 3** Feature extraction for subject 000, row: (**a**): 141 × 101 ear images, row (**b**): Convergence fields, row (**c**): Thresholded convergence maps.

images. The effect of brightness change by addition was also tested where we see that in the worst case where every odd image is subject to an addition of three standard deviations the force field results only change by 2%, whereas those for PCA under the same conditions fall by 36%, or by 16% for normalized intensity PCA, thus confirming that the technique is robust under variable lighting conditions.

These results are presented in Table 1 where Daugman's decidability index [12] combines the mean and standard deviation of the intra-class and inter-class measurement distributions giving a good single indication of the nature of the results. This index d' measures, how well separated the distributions are since recognition errors are caused by their overlap. The measure aims to give the highest scores to distributions with the widest separation between means, and smallest standard deviations. If the two means are μ_1 and μ_2 and the two standard deviations are σ_1 and σ_2 then d' is defined as

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}}. \quad (7)$$

Note that the best case index for PCA is slightly higher than the value of 3.43 obtained for the 141 × 101 images but this could be attributed to the reduction in data set size from 252 to 189 and also to the fact that the images have been extracted for PCA. Noise performance figures have also been included where noise has been modeled as additive noise with a zero mean Gaussian distribution. The signal to noise ratios of 6 dB, 12 dB, and 18 dB used are calculated as $20\log_{10}(S/N)$. We see that the technique enjoys excellent noise tolerance where even for an extreme noise ratio of 6 dB the performance only falls by about 3.6%. Interestingly, at a ratio of 18 dB

Physical Analogies for Ear Recognition. **Figure 4** First 4 eigenvectors for 111 × 73 pixel images.

the recognition rate actually improves over the noiseless recognition rate, but this must be put down to the combination of small changes and the random nature of the noise process. For reference, the corresponding noise results for PCA under the same conditions have also been included, where we see that PCA also performs well under noisy conditions but not quite as well as FFE at 6 dB where the fall is about 4.8%.

Summary

In the context of ear biometrics, a linear transform has been developed that transforms an ear image, with very powerful smoothing and without loss of information, into a smooth dome-shaped surface whose special shape facilitates a novel form of feature extraction that extracts the essential ear signature without the need for explicit ear extraction. It has been shown that the technique is robust under variable lighting conditions both by analysis and experiment. Convergence feature extraction has been described and it has been shown that it is a powerful extension to field line feature extraction. The technique has been validated by experiment where it has been shown that it compares favorably with PCA, especially under variable lighting conditions. In the process, a contribution has been made to the mounting evidence in support of the recognition potential of the human ear for biometrics.

Related Entries

- ▶ Ear Biometrics
- ▶ Earprints, Forensic Evidence of
- ▶ Holistic Ear Biometrics
- ▶ Overview of Ear Biometrics

References

1. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field energy functionals for image feature extraction. In: Proceedings of the 10th British Machine Vision Conference BMVC99, pp. 604–613. Nottingham, United Kingdom (1999)
2. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field energy functionals for image feature extraction. Image Vis. Comput. **20**, 311–317 (2002)
3. Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field feature extraction for ear biometrics. Comput. Vis. Image Underst. (CVIU) **98**(3), 491–512 (2005)
4. Hurley, D.J., Arbab-Zavar, B., Nixon, M.S.: Handbook of Biometrics, pp. 131–150. Springer, New York (2008)
5. Burge, M., Burger, W.: Ear biometrics in computer vision. In: Proceedings of the ICPR 2000, pp. 822–826. Barcelona, Spain (2002)
6. Yan, P., Bowyer, K.W.: Biometric recognition using three-dimensional ear shape. IEEE Trans. Pattern Anal. Mach. Intell. **29**(8), 1297–1308 (2007)
7. Moreno, B., Sanchez, A., Velez, J.F.: On the use of outer ear images for personal identification in security applications. In: Proceedings of the 33rd IEEE Annual International Carnahan Conference on Security Technology, 5–7 Oct 1999
8. Chen, H., Bhanu, B.: "3D free-form object recognition in range images using local surface patches". Pattern Recognit. Lett. **28**(10), 1252–1262 (2007)
9. Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G.: XM2VTSDB: the extended M2VTS database. In: Proceedings of the AVBPA'99, Washington, DC (1999)
10. Sadiku, M.N.O.: Elements of Electromagnetics, 2nd edn. Saunders College Publishing, Philadelphia, PA, (1989)
11. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. (March), **3**, 71–86 (1991)
12. Daugman, J.: Biometric decision landscapes. Technical Report TR482, University of Cambridge Computer Laboratory, Cambridge, UK (1999)

Physical and Logical Access Control convergence

The ability to authenticate and authorize a user across both physical and logical access control systems. The authentication may mean that a common credential, such as a biometric, is used across both systems, without the need for a user to re enroll. It may also mean that the user's authentication credentials are collectively stored in a mutually accessible repository that stores several biometric and nonbiometric types of authentication. The harmonization of the authorization of a user across physical and logical access control systems implies single enterprise policies, which are invoked in both systems.

► Access Control, Physical

Physics-Based Biometrics

► Biometric Sample Synthesis

Physics-Based Models

Physics-based models create mathematical equations that are derived from basic physical principles that can recreate biometric samples. An example of a physics-based model is a physics-based vocal tract model.

► Biometric Sample Synthesis

Piezoelectric

Piezoelectric is the property of crystals, such as quartz, which allows them to create electrical voltage when deformed under the application of a mechanical force.

► Digitizing Tablet

PIN Replacement

The method of using a biometric to unlock a smart card or SIM card instead of having to enter a PIN. The only secure method of PIN replacement is one that can perform the biometric authentication within the secure element itself, so that it remains in a locked state until user verification is complete.

► Transportable Asset Protection

Pitch

The pitch denotes the frequency at which the vocal cords open and close the larynx to produce the voiced sounds. The sound then produced resonates according to the shapes of the different cavities of the vocal apparatus. The pitch and resonances are generally measured through a frequency based analysis of the speech signal to extract information dependent to the speaker.

► Transportable Asset Protection

Pixel

Pixel is a single picture element of the image. It never corresponds to the physical pixel on the camera sensor. The size of the pixel depends on the image (and the camera) aspect ratio (the ratio between the image width and length).

► Biometric Sensor and Device, Overview
► Image Feature Extraction
► Image Formation

Platen

In a fingerprint device, the term platen refers to the surface of a sensor on which the finger should roll or

adhere. It can be done of different materials, but it is mainly done with glass. In a hand-geometry device, it means the flat surface of the hand geometry reader on which the person presented to the device places his or her hand. The platen is usually equipped with a number of pegs (or pins) to guide the placement of the person's hand and to ensure the accurate measuring of the hand's geometric structure.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor
► Hand-Geometry Device

Point-Light Display

The motion of a living thing depicted by just the motion of specific points on the body. Originally generated by filming actors in darkened rooms, wearing dark clothes with lights attached to the joints. By adjusting the contrast of these films only the motion of the lights was visible. Modern day point-light displays are generally created by movement data obtained by three-dimensional motion capture equipment such as the one used for computer animation. The point-light display shows the motion of the actor with the majority of other person information subtracted from the display.

► Psychology of Gait and Action Recognition

Polar

Polar images are recorded as a matrix in which each element of the matrix is an intensity value whose location is expressed in polar coordinates. The origin of the coordinate system is defined as the center of the image, and the location of each intensity sample is expressed as a radial distance r from the center along a particular angular direction θ. Typically, each row in the image matrix contains all of the angular samples for a particular radial distance, and each column of the image matrix contains all of the radial samples for a particular angular direction.

► Iris Image Data Interchange Formats, Standardization

Polarized

Of or relating to one or more poles (as of a spherical body).

► Iris Standards Progression

Police Law Enforcement

► Law Enforcement

Portal

► Iris on the Move

Pose

The angle of the head relative to the camera-to-subject line.

► Photography for Face Image Data

Pose and Motion Models

Models that describe what combinations of joint angles are plausible and how they can vary over time in typical human motions. They are often expressed probabilistically and specify the probability of a single 3D pose or of a set of consecutive poses.

► Markerless 3D Human Motion Capture from Images

Poststratification

Poststratification is a statistical technique that forms strata of observations after the data has been collected to better inform statistical inference.

▶ Test Sample and Size

Potential Energy Transform

An invertible linear transform which transforms an image into an energy field by treating the pixels as an array of particles that act as the source of a Gaussian potential energy field. It is assumed that there is a spherically symmetrical potential energy field generated by each pixel, so that $E(\mathbf{r}_j)$ is the total potential energy imparted to a pixel of unit intensity at the pixel location with position vector $\mathbf{r}_j$ by the energy fields of remote pixels with position vectors $\mathbf{r}_i$ and pixel intensities $P(\mathbf{r}_i)$, and is given by the scalar summation,

$$E(\mathbf{r}_j) = \sum_i \left\{ \begin{array}{l} \dfrac{P(\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|} \forall i \neq j \\ 0 \forall i = j \end{array} \right\}. \qquad (1)$$

To calculate the energy field for the entire image, Eq. 10 should be applied at every pixel position. For efficiency this is calculated in the frequency domain using Eq. 11 where $\Im$ stands for FFT and $\Im^{-1}$ stands for inverse FFT.

$$\text{energyfield} = \sqrt{M \times N}\Im^{-1}[\Im(\text{unit energy field}) \times \Im(\text{image})]. \qquad (2)$$

▶ Physical Analogies for Ear Recognition

Preemployment Screening

▶ Background Checks

Pre-Processing

Palmprint pre-processing refers generally to the extraction of the region of interest and its normalization. A coordinate system is defined to align different palmprint images for matching. Normally, the central part of a palmprint is extracted from the image boundaries for reliable feature measurements. The extracted central part is further subjected to histogram equalization.

▶ Palmprint Features

Pretty Good Privacy (PGP)

Pretty Good Privacy (PGP) is a computer program that provides cryptographic privacy and authentication. It was originally created by Philip Zimmermann in 1991.

▶ Fingerprints Hashing

Primary Biometric Identifier

Anatomical and behavioral characteristics such as fingerprint, palmprint, face, iris, hand-geometry, voice, signature, and gait can be used to reliably determine or verify a person's identity. These biometric traits constitute a strong and permanent "link" between a person and his identity and these traits cannot be easily lost or forgotten or shared or forged. Hence, they are known as primary biometric identifiers and the systems that recognize people based on such traits are referred to as primary biometric systems.

▶ Soft Biometrics

Principal Component Analysis

See PCA (Principal Component Analysis)

Principal Curvatures

The maximum and minimum values of the normal curvature at a point on the surface are called the principal curvatures. At any given point on the surface, intersection of the surface with a plane containing the normal vector and a particular tangent direction forms a curve. Curvature of this curve is called the normal curvature. Curvature values in all tangent directions form the normal curvatures at that point.

► Palmprint Features

Principal Lines

Human palmprints show a number of well defined lines. These include major palmlines, often called the principal lines, and wrinkles. Principal lines and wrinkles on palm are distinguished by their position and thickness. Most palmprints show three principal lines: heart line, head line and life line. Properties of the principal lines can be summarized as follows:

1. Each principal line meets the side of the palm at approximate right angle, when it flows out
2. The life line is located at the inside part of the palm, which gradually inclines to the inside of the palm
3. In most cases, the life line and head line flow out of the palm at the same point

► Palmprint Features

Privacy

Privacy is the ability or right of an individual to control how information pertaining to that person is collected, distributed, and used. Surveillance, by its premise violates or at the least infringes on the privacy of an individual. Hence, in that regard, it is extremely important to ensure that the information collected (with or without consent) does not violate the legal rights of the individual.

► Surveillance

Privacy Issues

Terence M. Sim
School of Computing, National University of Singapore, Singapore

Synonym

Data protection

Definition

Privacy is a multidimensional and evolving concept, whose definition varies according to country and culture. It is thus difficult to agree on a precise definition that is universally accepted. Nevertheless, certain notions of privacy have become fairly standard, especially among industrialized nations. These include: informational, bodily, territorial, and communications privacy. Of these, biometrics not only impacts informational privacy, but also affects bodily and territorial privacy as well. However, contrary to the claims of civil libertarians and to Hollywood hype, biometrics need not be antithetical to privacy. Indeed, by understanding the relevant issues, it is possible to design into a biometrics system measures that will safeguard, and even enhance, privacy. Doing so will increase user acceptance of biometric systems, or at least, render such deployments more tolerable.

Introduction

Losing one's privacy often entails consequences. At best, the consequence is fortuitous, as when receiving an unsolicited discount on a product that one was about to purchase. At worst, loss of privacy could result in harassment, marital discord, or even termination of employment. Biometrics, by its very nature, impacts privacy. Civil liberty advocates routinely decry its use,

and Hollywood movies often portray it negatively. Yet biometrics, when judiciously used, can in fact enhance privacy by improving security to sensitive data. This chapter examines the privacy issues arising from the deployment of biometric systems, and suggests ways in which careful system design, along with government policies, industry best practices, research and education, can ameliorate privacy concerns, leading to greater user acceptance of such systems.

Privacy is a relatively new concept, dating back to around 1900 A.D. Although some authors have tried to argue that privacy notions may be found in ancient religious texts such as the Christian Bible and the Islamic Koran, most would agree that privacy appeared in public consciousness and began influencing public policy only at the turn of the last century. In Europe, privacy notions came about largely in response to rapid urbanization and the horrors of two World Wars, leading to its formal articulation in the 1950 European Convention on Human Rights (ECHR). In the U.S., the 1890 journal article by Warren and Brandeis [1] may be considered a defining moment.

In other parts of the world, the development of privacy often mirrors economic and political progress. Privacy laws are most developed in nations that have achieved high economic output, and where democratic principles have been established for some time. Privacy issues are by no means static; new technologies often reveal subtle nuances in prevailing notions, and expose inadequacies in existing laws, leading to new regulations or amendments being proposed. Changing user perceptions, brought about by increasing technological savvy, also play a role in shaping public policy on privacy. An historical account of privacy developments in different countries, no doubt an interesting study, is outside the scope of this chapter. For general privacy issues, please see [2–5].

Privacy Notions

Although actual definitions vary depending on culture and country, four notions of privacy are almost universally accepted:

- *Informational privacy:* This pertains to the collection and subsequent usage of ► personal data.
- *Bodily privacy:* This concerns protecting the physical body from invasive procedures.
- *Territorial privacy:* This relates to the observation of one's activities in a physical space, especially one's home or bedroom, but also in public places where anonymity is to be expected.
- *Communications privacy:* This addresses the protection of one's communications, such as emails, letters, and telephone conversations.

Informational privacy can in turn be understood in terms of four concepts: ► unnecessary data collection, unauthorized data collection, ► unauthorized data disclosure, and ► function creep.

By definition, biometrics *is* personal data: a biometric sample is a measurement of a human body for the purpose of identifying that person. Moreover, in typical usage, other personal data, e.g. date of birth or address, are also retrieved along with the identification. Clearly then, biometrics impacts informational privacy. But there are other concerns as well. For some biometrics, notably DNA samples and retina scans, the very act of acquiring the biometric may be considered an invasion into one's body. These types of biometrics, thus, affect bodily privacy. Likewise, for end-users who regard physical contact as unhygienic, placing one's finger on a fingerprint sensor may constitute a violation of bodily privacy. Such an aversion to physical contact will be especially acute during an epidemic, as when SARS broke out in 2003 A.D. Finally, there may be issues with territorial privacy as well. Whenever face recognition is coupled with surveillance cameras to monitor a public space, one's anonymity is lost when traveling through that space. Biometrics that can be remotely acquired without the user's cooperation, such as face images, and to a lesser extent voice patterns, potentially increase territorial privacy risks. Unfortunately, with expected improvements in sensor technology, more types of biometrics will be amenable to remote acquisition, even for those that currently require physical contact.

Biometrics is inherently privacy-neutral: it neither enhances nor enervates privacy. It is no different from a database record of a person's particulars. The real concern is how biometrics is used, or more precisely, misused. The misuse of biometrics is more potentially more damaging than the misuse of a database record because one cannot lie about biometrics the way one can falsify one's name or address. For example, one cannot give a fake photograph when pressured by a salesperson to sign up for some dubious product offer. Moreover, biometrics is generally permanent

throughout a person's lifetime, and thus cannot be revoked once compromised (unlike changing a password, PIN, or even one's name). This immunity from falsification and revocation makes biometrics a good choice as an universal identifier. For example, banks, government agencies, and supermarkets may use the thumb print for verification. The convenience of using a single fingerprint to access one's bank account, to obtain government services, and to pay for groceries is extremely compelling for end-users and organizations alike. But so are the risks correspondingly magnified. The linking of the bank's database with those of the government and the supermarket to monitor one's intimate details would, in most places, be considered an egregious invasion of privacy. Even if such monitoring were sanctioned by the appropriate authority, the victim is unlikely to derive much comfort.

Privacy is not the same as security. Privacy is concerned with *people* (their intimate details, personal space), whereas security has to do with *systems*. A good privacy policy protects people, whereas a secure system is one that is effective at preventing unauthorized access to the resource being protected. Identity theft (the use of someone else's identity for personal gain) is primarily a security breach, but also a privacy violation. Thus, privacy begins with securing the biometric system itself. An insecure biometric system affords little privacy protection.

Effectively addressing the privacy issues arising from the deployment of biometric systems require a holistic and multipronged approach. This chapter highlights five ways, as described in the following sections.

System Issues

Privacy should not be an afterthought; rather, measures to safeguard privacy should be designed into a biometric system right from the beginning. Good strategies for doing this may be found in [6, 7]. In addition, [8] has two chapters on the privacy aspects of biometrics in relation to U.S. and European laws. The following highlights the main issues.

1. *Alternative technologies.* Consider nonbiometric alternatives. Biometrics is not the only technology for verifying identity or controlling access to a protected resource. Other technologies may be more appropriate, and less privacy invasive. For instance, the humble lock-and-key works very well for gymnasium lockers, and replacing it with a fingerprint access control system seems excessive. Besides the obvious privacy concerns, the fingerprint system does not permit the ad hoc transfer of authorization, as when asking a friend to retrieve one's belongings from the locker. The low-tech lock-and-key has no such problem.
2. *Choice of biometrics.* Choose a biometric appropriate for the application, taking into account cultural and religious sensitivities. As mentioned before, DNA and retina scans may be regarded as invading one's bodily privacy because of the way these samples are collected. Since DNA can reveal genetic defects and retina scans can reveal diabetes, their usage can lead to function creep (also see below). Likewise, avoid choosing biometrics that requires physical contact for acquisition when doing so would alarm end-users who consider such contact unhygienic. Finally, using face recognition may offend the modesty and privacy of end-users who veil their faces for religious reasons.
3. *Template storage.* To enhance privacy, templates must be securely stored, preferably with a strong encryption method. Moreover, distributed storage is preferred over a centralized database. Where possible, delete the template as soon as it is no longer required. This is preferable to storing it indefinitely. Finally, allowing the end-user to decide when and how the template can be used reduces privacy risks. This could be achieved by permitting the end-user to opt in or out (of using the biometric system) at the end-user's discretion, or to specify the encryption method, or the duration and location of template storage. In this regard, the ▶ Match-on-Card technology for fingerprint verification, in which all the steps in the biometrics architecture are implemented on the smart card itself, comes closest to fulfilling this privacy ideal.
4. *Function creep.* Once a biometric system is operational, it is often convenient to use it for other purposes. From a privacy perspective, this must be resisted, even if the secondary purpose is a noble one. At the very least, consent must be obtained from the end-user for the expanded scope of biometric usage. This is especially true for biometrics that reveals more than just identity, e.g., DNA and retina scans that reveal medical conditions, finger vein patterns that reveal blood oxygen level, and face images that reveal gender, ethnicity, and

approximate age. The potential for medical, sexual, racial, and age-related discrimination arising from using such nonidentity information is clear. At times, function creep can be subtle, as the next example illustrates: To improve security, a secondary school (name suppressed to avoid embarrassment) implemented a fingerprint system to control access to its science and computer laboratories. It soon discovered a serendipitous way to reduce its electricity bill: by identifying persons leaving the labs without switching off the air conditioner.

Government

Governments play three important roles in ensuring that biometric usage protects privacy: by enacting appropriate laws, by self-regulation, and by cooperating with other governments. Privacy requires legal backing to be effective. Yet laws are notoriously difficult to get passed. A quick survey of the state of privacy laws in selected countries is shown in Table 1. It is clear from the table that constitutional provision for privacy does not automatically guarantee stronger privacy laws. For example, both Canada and Australia do not provide for privacy rights in their Constitutions, yet both have enacted a Privacy Act and a Privacy Commissioner to oversee and prosecute privacy violations. China and the U.S. are opposite examples, having some form of privacy rights in their Constitutions but no specific Privacy Act. Instead, both rely on a hodgepodge of sectoral laws to regulate privacy. Besides having the right laws, it is also necessary to review them periodically as biometrics may engender unanticipated privacy issues. Nevertheless, countries having an explicit Privacy Act and appointing a Privacy Commissioner are expressing a strong commitment to protect privacy. The experience of Australia, Canada, and Hong Kong are thus worth learning from.

Privacy Issues. **Table 1** Privacy laws in selected countries, excerpted from [2, 3]

Country/region	Privacy in constitution?	Related laws	Privacy act?	Privacy commissioner?
Australia	×	Crimes Act, Telecommunications Act, Data-Matching Program Act	Privacy Act 1988	✓
Canada	×	PIPEDA, Telecommunications Act, Bank Act	Privacy Act 1983	✓
China	Limited rights	Civil Law, Practice Physician Law, Law on Lawyers	×	×
European Union	1950 European Convention on Human Rights	Telecommunications Privacy Directive	Personal Data Directive 1995	×
Hong Kong S.A.R.	in Basic Law	—	Personal Data (Privacy) Ordinance 1996	✓
Japan	Article 13	—	Protection of Personal Information Act 2003	various Ministers
Singapore	×	Banking Act, Computer Misuse Act	×	×
United States	Not explicit	Video Privacy Protection Act, Right to Financial Privacy Act	Privacy Act 1974 (limited to govt)	×

In most countries, the Government is also the largest deployer of biometric systems. Thus, governments can strongly influence how biometrics is used by practicing self-regulation, and maintaining transparency and accountability. Governments should not, for example, covertly deploy surveillance systems, nor share sensitive data between agencies across different jurisdictions. Federal and local authorities should also respect each other's boundaries. These prescriptions are self-evident, perhaps even naïve. Alas, they are usually circumvented in the name of national security and expediency. Most nations have emergency laws that can be invoked when confronting terrorism threats or disease epidemics. Such laws usually give the government carte blanche power, to the detriment of privacy concerns. While citizens are generally willing to give up some privacy in exchange for personal safety during times of threat, governments are less willing to relinquish their power once the crisis subsides. This asymmetry ought not to exist, and should be corrected.

The third important role of the government is international cooperation. In this digital age of transborder data flows, privacy is only as strong as the weakest jurisdiction. Already, regional and international standards have been drawn up to address common privacy issues. Examples include the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, the European Union Personal Data Directive, and the Asia-Pacific Economic Cooperation Privacy Framework. Arguably, more can be done to ratify and implement these guidelines in a timely manner across member countries.

Industry

Privacy is too important to be left in the hands of governments. Complementing government legislation, and often faster to enact, is a necessary set of industry regulations developed and updated by a nonpartisan biometrics association. Such an association should ideally consist of vendors, end-users, legal counselors, and academics. Its role is to promote best practices, certify compliance of vendor products with international biometrics standards, educate the public, and otherwise regulate the industry.

An example of this is the nonprofit Biometrics Institute in Australia [9], which recently drafted a Privacy Code and obtained the approval of the country's Office of the Privacy Commissioner (OPC). The Code, essentially an industry-specific realization of Australia's National Privacy Principles, recommends guidelines for how biometrics data should be collected, used, and disclosed, among other things, to safeguard privacy. Members of the Institute voluntarily subscribe to the Code, thereby agreeing to be bound by it. In return, the subscriber establishes itself as a trustworthy party, and gains exclusive rights to bid for government projects that require privacy certification. Code violations are handled directly by the OPC, which, unlike the Institute, has the legal teeth to prosecute. This symbiotic relationship between the Institute and the OPC is admirable, and greatly enhances public trust. Independent of the Code, the Institute also conducts privacy impact assessments and educational talks for members.

Another essential role of the industry is to track and participate in international standards, usually in partnership with a government standards body. The ISO JTC 1/SC 37 is the Biometrics Technical Subcommittee under the ISO umbrella [10]. Its biannual meetings divide into several workgroups, the sixth of which concerns the "Cross-Jurisdictional and Societal Aspects of Biometrics," thus encompassing privacy issues. Of interest is the still under development technical reference ISO/IEC DTR 24714 parts 1 and 2, which lists 15 privacy principles for biometric systems. Although not yet publicly available, these documents are useful for reference and for adapting to suit country-specific norms.

Research

Since biometrics is a technology, it seems plausible to combat its "evils" with more technology. So-called Privacy Enhancing Technologies (PET) aim at protecting privacy while enabling their benefits to be enjoyed. One possibility centers around the problem of authorization without identification.

For many applications, it is not the identity of the end-user that matters, but only whether the end-user is duly authorized. That is, what needs to be established is a *proof of authorization* (or proof of credentials) rather than a proof of identity. This is the case for all access control systems (Is the end-user authorized to gain access to the protected resource?),

and subscription systems (Does the end-user have the necessary credentials for this service?). Proof of authorization is a weaker notion than proof of identity, and can in fact be achieved using a cryptographic technique called *zero-knowledge proof*. Related research includes *group signatures* and *k-anonymity*, both of which permit identification only of a group of people, but not the individual within it. This coarser form of identification is like protecting a room with a traditional lock-and-key, and giving the keys only to the group of authorized people. In effect, the key authorizes the group while preserving individual anonymity. For more details, please see [11, 12].

A slight generalization of this concept is the *proof of authorized role*. Here, the same person may assume different roles when interacting with a system. For example, it is common practice for a person to login to a computer system either as an administrator, or as a normal user, depending on the purpose of usage. The required proof is not so much identity, but which role the person wishes to assume. Different biometrics may be associated with each authorized role to facilitate interaction with the system.

There is yet another notion of identity: for certain applications, not only is the identity required, but the physical presence of the end-user must be guaranteed. This *proof of presence* is clearly a stronger requirement. An example of this is during wedding ceremonies (be they religious or civil), where additional witnesses are usually called upon to prove the identities and presence of the marrying couple. Another example is at the polling station, where it is necessary to establish that the voter is physically present to cast his or her vote, instead of relying on a proxy.

Distinguishing between these subtle notions of identity is important. An authentication based on biometrics is really a proof of presence, because the biometric sample is collected "live" from the person. Thus, using biometrics in situations that require only a proof of authorization may be an overkill, and can lead to privacy abuse. Research in PET is still in its early stages, but should be encouraged and funded.

Education

The cliché, "perception is reality," is especially true for biometrics, where misconceptions and hyperbole abound. Hollywood movies such as *Minority Report (2002)* and *Gattaca (1997)* tend to negatively portray biometrics as powerful tools used by Big Brother regimes to track individuals. Media reports of high profile abuses involving biometrics further fuel public mistrust. Occasionally, even well-meaning privacy advocates unwittingly deride biometrics more than the technology deserves. Such poor perceptions can lead to public resistance, and even sabotage, of biometric systems.

It is, therefore, important to increase public awareness through educational talks and open dialog among vendors, deployers, end-users, and privacy advocates. Besides the technological issues, privacy issues have to be realistically addressed, including assuring the end-user on what recourse is available should he or she feel victimized. Such educational talks can be organized by anyone, although it is better received if it comes from a neutral party, such as the non-partisan biometrics association (see above). Dialog should also be on-going, because new issues are constantly thrown up.

Future Concerns

Many advocates believe that privacy is increasingly under attack from two main fronts. New technologies that permit more efficient data sharing, or that facilitate covert surveillance or identification of individuals, pose real threats. Also, the rise of terrorism and the imminence of epidemics (such as avian flu) necessitate governments to more closely monitor the movements and activities of their citizens in a bid to, ironically, protect the same citizens.

Biometrics technology will also be further developed. Besides providing proof of authorization or identity, biometrics may soon be able to reveal emotional states. It is already possible to detect anxiety in the voice, adding a new level of privacy concern to voice (speaker) recognition. Other emotions may yet be detectable through current or novel biometrics, and could lead to a revolutionary type of lie detector.

The emerging field of *neuroeconomics* [13] attempts to understand brain activity (measured through functional magnetic resonance imaging, or fMRI) and economic decisions such as buying a product. Researchers are able to predict, from fMRI patterns, whether or not a person is about to make a purchase. Other research in

the field of cognitive science have demonstrated that fMRI scans can detect cognitive states in the brain. From such developments, it is but a small step to imagine the day that biometrics (in the broader sense of "bodily measurements") will offer *proof of intention*, i.e. a kind of mind reading. The privacy abuses arising from such a clarivoyant technology are too frightening to even contemplate.

In light of all these, what can be said about the future of privacy? One line of argument, proffered by James Rule [5], is that privacy goals are still eminently feasible. To quote:

> The issues involved are ultimately ethical and political, not technological. If we determine to do so, we can readily implement systems that place the burden of justification on those who would create personal data systems in the first place, ..., that limit the amount and variety of personal data allowed to bear on determinations of how organizations will treat individuals.

However, doing this requires accepting the social cost that information systems will necessarily be less efficient because of increased privacy checks. Current systems are predicated on providing better services or making better decisions through gathering more personal data. Only by abandoning this avarice for efficiency can society hope to restore privacy.

The other school of thought takes the opposite view: that privacy is merely a passing ideal of the previous century, increasingly irrelevant for the twenty-first century. There will be no privacy in the future, and the sooner we get used to it, the better. Among such proponents is Scott McNealy, Chairman of Sun Microsystems, who famously quipped that "privacy is dead." Increasingly, young people today act as if this is true [14]. They are not afraid to reveal intimate details in social networking sites, or post in their blogs videos of themselves in situations deemed highly embarrassing just one generation ago. They do this despite knowing the privacy risks. They accept that their daily activities, social habits, and personal data can be viewed by anyone. Yet life goes on. To be sure, much private data have only temporary value. For instance, one's soda and sartorial tastes are ephemeral, valid only until the next fashion wind blows. For these young people, losing one's privacy in such matters is hardly worth losing sleep over.

Related Entries

- ▶ Biometrics Architecture
- ▶ Match-on-Card
- ▶ Security

References

1. Warren, S., Brandeis, L.: The right to privacy. Harvard Law Rev. **IV**(5) (1890)
2. Electronic Privacy Information Center: (1994). http://www.epic.org
3. Privacy International: (1990). http://www.privacyinternational.org
4. Staples, W.G. (ed.): Encyclopedia of Privacy. Greenwood, Connecticut (2007)
5. Rule, J.B.: Privacy in Peril. Oxford University Press, New York (2007)
6. National Science and Technology Council: Privacy and Biometrics, Building a Conceptual Foundation (2006). http://www.biometrics.gov/docs/privacy.pdf
7. Nanavati, S., Thieme, M., Nanavati, R.: Biometrics – Identity Verification in a Networked World. Wiley, New York (2002)
8. Wayman, J.L., Jain, A.K., Maltoni, D., Maio, D. (eds.): Biometric Systems: Technology, Design and Performance Evaluation. Springer, New York (2005)
9. Biometrics Institute Ltd.: (2001). http://www.biometricsinstitute.org
10. ISO JTC 1/SC 37: International Organization for Standardization Biometrics Technical Subcommittee (2002). http://www.iso.org/iso/home.htm
11. Camenisch, J., Lysyanskaya, A.: An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In: Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques, pp. 93–118. Innsburck, Austria (2001)
12. Chaum, D., van Heyst, E.: Group signatures. In: Advances in Cryptology, pp. 257–265. Brighton, UK (1991)
13. George Loewenstein and Scott Rick and Jonathan D. Cohen: Neuroeconomics. Ann. Rev. Psychol. **59** (2008)
14. Nussbaum, E.: Say Everything (2007). http://nymag.com/news/features/27341/

PrivateID™

PrivateID™ is an image processing protocol and data standard that enables a Proof Positive-certified iris

camera to capture an image, process it and prepare it for transport in the most secure way possible.

► Iris Acquisition Device

Privium

► Iris Recognition at Airports and Border-Crossings
► Simplifying Passenger Travel Program

Probabilty Density Function (PDF)

The statistical function that shows how the density of possible observations in a population is distributed.

► Gaussian Mixture Models

Process Artifacts

When footwear outsole patterns are created either through pressing or molding, certain defect in the pressing tools and mould will contribute to artifacts being left on the final product. A common artifact formed in conjunction with the molding process is due to tiny air bubbles being trapped within the mould, leaving gaps in the outsole pattern.

► Footwear Recognition

Procrustes Shape Distance

The Procrustes shape distance is a metric that captures the shape of an object independent of the set of Euclidian transformations (translation, rotation, and scale). The Procrustes distance computation assumes that all objects can be represented by a set of landmark points, each object has the same number of points, and exact correspondence between the points is known from one object to the next.

► Hand Shape

Proposal Descriptive and Decision Making Model

P.J. van Koppen (University of Leiden and University of Antwerp), and H.F.M. Crombag, (University of Maastricht) analyzed all types of forensic evidence and formulated the common, basic requirements in an article published in the *Dutch Journal for Lawyers* in January 2000. These are as follows:

1. The expert has a descriptive model at his disposal that describes the relevant characteristics for comparison and identification of the mark found at the crime scene, with the characteristics of the defendant.
2. There is sufficient variation between different persons regarding these relevant characteristics.
3. The relevant characteristics change very little over time that even after some time comparison is feasible.
4. The expert has a method with which the relevant characteristics can be established unequivocally/unmistakably.
5. The expert has rules of decision-making at his disposal with the help of which he can decide about identification, based upon the comparison.

► Fingerprint Matching, Manual

Prosody

Prosody concerns the "melody" of an utterance. As such, prosodic aspects of a sentence are rhythm, intonation, and stress/emphasis. Acoustical expressions of prosody are duration (of syllables/phonemes), loudness, pitch and even formant structure (which might be different in stressed vowels than in unstressed vowels).

► Voice Sample Synthesis

Protocol

Protocol is the preset procedure by which each template in the database is compared to some specific feature during the performance evaluation of the fingerprint-matching approach, and the rule to treat the failed matching.

► Fingerprint Matching, Automatic

Pseudo-Random Number Generator

An algorithm to generate a sequence of numbers that approximate the properties of random numbers. The sequence is not truly random since it is completely determined by a few parameters and initial values. Pseudo-random numbers are useful for cryptography and simulation.

► SFinGe

Psychology of Gait and Action Recognition

Frank E. Pollick
Department of Psychology, University of Glasgow, Glasgow, UK

Synonyms

Action categorization; Action understanding; Biological motion perception

Definition

The psychology of gait and action recognition strives to understand the processes that underlie how people detect, recognize and understand the movements of others. Since gait is a fundamental human activity, it has formed an important visual signal for psychologists to examine. Experiments have shown that sparse representations of gait support the recognition of identity, gender, and emotion by observers even when viewing conditions are degraded. The study of gait and action recognition focuses on several questions, including: what visual properties uniquely specify human movement; how to quantify human performance in action recognition; and the neural mechanisms that form the basis of decoding human movement.

Introduction

The modern study of the *psychology of human movement*, in particular the perception of gait, starts with the work of the Swedish Psychologist Gunnar Johansson in the 1970s [1]. The work of Johansson and his contemporaries focused on how humans use motion to infer the structure of objects moving in the world. To demonstrate this capability, he attached lights to the joints (elbow, shoulder, ankle, etc.) of an actor and filmed the actor moving about in a darkened room. In any individual frame of the movie, the points did not convey a strong impression of structure. However, when the movie was played, a vivid impression of the actor moving through space was obtained (Fig. 1). These displays of human activity are called ► point-light displays and the general field of studying how the individual point motions spontaneously organize into the percept of a moving body is known as ► biological motion perception [2].

There are several reasons why point-light displays form a key contribution to the psychology of understanding human actions. The first is that point-light displays represent an action as just the 2D locations of a set of joint locations on the body and thereby remove a multitude of other visual information that can be conveyed by things like hairstyle, clothes, facial expression, and other factors; thus the contribution of motion itself can be effectively isolated. A second reason is that the relevance of motion is highlighted, since for these displays any particular static frame typically does not elicit a strong impression of a body. A third reason is that the motion properties of a small set of points can be easily quantified, allowing for biological motion displays to be compared in experiments against other motion patterns with identical motion

Psychology of Gait and Action Recognition. **Figure 1** Examples of frames taken from an image sequence of a point-light display of a ballet dancer. Each individual frame is seen predominately as just a collection of points, although a static human form is possibly visible in some frames. However, a vivid impression of the action is appreciated nearly instantaneously when all the frames of the sequence are presented consecutively in a motion sequence.

statistics. Finally, point-light displays of biological motion provide a compelling demonstration of the use of motion to perceptually organize a display, the precise mechanisms of which are still unclear.

One issue with the use of point-light displays is how to evaluate the role of form perception. It could be assumed that the mechanisms behind the perceptual organization of the point-light displays provide ► structure-from-motion information regarding form [3]. However, this form information is not necessarily equivalent to what might be available from that presented directly from an image or even some other reduced form such as a stick figure or a silhouette. The distinction between form and motion is important from the psychological perspective since data from neuroscience support the idea that the human visual system is segregated into largely distinct pathways that specialize in processing form and motion information.

The study of point-light displays has been critical in developing the understanding of how motion can be used to convey the presence of an actor from minimal information, and to a degree psychological research has focused in this domain. However, one other question which has drawn attention is just what person properties can be derived from the point-light displays. Namely, can actor qualities such as identity, gender, emotion, attractiveness, and intent be identified from such displays? Experimental results generally indicate that human observers can identify such qualities at better than chance, though what cues they use and how to evaluate their performance on an absolute scale are presently areas of active research.

Psychological Studies into Perceiving Biological Motion and Recognizing Person Properties

In this section empirical investigations into biological motion perception and action recognition is reviewed. The majority of research into biological motion perception has involved point-light displays of gait. Research into action recognition has typically also used point-light displays, though sometimes limited the visual display to those points that change substantially for the different actions. Another methodological difference between studies of biological motion perception and action recognition has been that research into biological motion perception has typically relied upon psychophysical analyses of the ability of observers, under normal and degraded viewing conditions, to detect the presence of a walker or to discriminate the walking direction [2]. In contrast, research into action recognition has used a variety of experimental techniques aimed at uncovering the underlying features used by observers to recognize the action being performed as well as properties of the actor.

Numerous experiments have shown that the ability of the perceptual system to detect the biological motion of a walker is surprisingly resistant to distortions of the walker or the embedding of the walker in visual noise. For instance, limiting the lifetime of the points on the walker or displacing them to points on the skeleton, as opposed to joint locations, barely diminishes the ability of an observer to detect a point light walker. Furthermore, masking the motion of the points using a background of randomly moving noise dots still does not greatly reduce the impression of a human walker unless the masking noise is used in

combination with disruptions to the synchronization between points on the walker [4]. This disruption to synchronization, by introducing time delays and advances among the points of the walker, renders local motion cues ineffective for detecting biological motion or discriminating the direction of motion, forcing the perceiver to rely on global or configural cues [5]. One effective way to mask the motion of a walker is to take the walker points themselves as the source of points to use as background noise. However, even when this is done and the masking dots contain local motion signals identical to those of the walker, large numbers are required to diminish the impression of a human walker [6].

The apparent fine tuning of the perceptual system to point-light displays of walking has raised the question of whether or not specialized motion detectors are involved in the processing of biological motion [7]. Research comparing the perception of biological motion to other kinds of motion has revealed differences in motion tuning characteristics. Namely, that the processing of biological motion involves the integration of motion information over a larger spatial extent and a longer temporal window than that found for other types of motion [8]. However, these results fall short of proving that dedicated biological motion detectors exist since they reflect the output of the entire action processing system, which might include specialized higher-order mechanisms for processing human actions that augment standard motion detectors. Evidence for the involvement of higher level factors comes from the breakdown of biological motion detection when the local form and motion relations are preserved, but the entire display is inverted [9]. Perception of these inverted point-light walkers is impaired relative to upright walkers, independent of the location of the source of gravity [10].

Investigations of action recognition from point-light displays have shown that a variety of actor properties and action styles can be reported above chance [2, 11]. Importantly, for the field of biometrics, it has been shown that human observers can recognize identity from point-light displays of gait [12, 13, 14, 15]. The work of Stevenage [12] also compared recognition of identity from point-light displays to video recordings of the same actors under full light and diminished light conditions. It was found that identification performance was equivalent between the different viewing conditions and this was taken as strong evidence that the motion cues contained in gait were sufficient to provide cues to identity. Further evidence for the importance of motion cues comes from results which show that even when size and walking frequency are made equal for all the targets to be recognized, performance decreases but recognition of identity is still greater than chance and generalizes to novel viewing directions [16].

Other actor properties which can be recognized from point-light displays of gait include gender, emotion and even vulnerability [17]. Emotion can also be obtained from point-light displays of whole-body dance movements as well as just the arm performing everyday movements such as knocking. In sports a variety of athlete characteristics and movement intentions can be gleaned from observing the action [18]. As might be expected from the variety of scenarios discussed, there is not a specific single action feature that has been found to explain the recognition of actor properties and action styles. However, researchers have generally distinguished between form and motion cues. For example, in gender recognition, experiments have focused on the diagnosticity of form cues encoded in the different relative sizes of hips and shoulders, while other studies have concentrated on differences of hip motion [19]. In general, both form and motion features appear effective to inform recognition and given the complexity of human motion it is hard to tease apart the different sources even when using point-light displays.

Even with simple actions and extremely reduced point-light displays there is a complex pattern of body postures that unfold in time, and it is an open question as to what features within this signal are crucial. One way this complexity has been addressed in cases such as gender recognition has been to reason from first principles about what features drive recognition, and to use carefully manipulated action displays to test hypotheses about these features. However, another approach has been to use techniques of automatic *pattern classification* to quantify how information in the point-light displays are used for recognition. An issue with automatic *pattern classifiers* is that while they can effectively categorize action styles they do not necessarily provide intuition into what specific features differentiate the styles. For this reason they have been applied in two domains that do not require an intuitive understanding of the features. One of these is to quantify levels of human performance and the other is to invert the *pattern classifiers* so that

recognizable differences in action style can be injected into movements.

The use of *pattern classifiers* to quantify human performance in recognition has been achieved by using classifiers as the standard of comparison for human performance. If the classifier can be shown to optimally achieve recognition by using all the available information then the efficiency of human performance can be expressed as the percentage of available information used by the human observers. If, however, it cannot be shown that the classifier optimally uses all the information available then it is still possible to use the classifier to estimate an upper bound of human efficiency or to compare human recognition of different action properties against a standard classifier [20]. For the case of recognizing gender from point-light displays of gait it has been shown that the average percentage correct in gender identification is 66%, which is moderately above chance of 50%. Efficiency at gender recognition, calculated relative to a model emphasizing structural features [21] is 26%, which is a high value since efficiency values of 10% or higher are generally considered excellent performance. This low percentage and high efficiency reflects either that the male and female distributions are highly overlapping and that humans are very effective in using the available structural information, or that since the structural features do not incorporate motion information that the efficiency results are inflated [22]. In summary, the calculation of efficiency provides a valuable tool to examine the recognition of human movement and provides a means to use methods of automatic *pattern classification* or to examine how performance relates to the modeled use of a specific feature.

Another application of automatic *pattern classifiers* has been to "invert" their performance so that instead of recognizing actions they are injecting style into normal movements or otherwise modifying the movements [23, 24]. The intuition behind this is that the action of a point-light display can be specified by the three-dimensional coordinates of the body sampled many times per second, resulting in thousands of values representing even a simple action. Each action can be considered as a point in this high dimensional space and the different styles of the action as different regions of this movement space. By obtaining classifiers to identify these different regions, possibly with the use of *dimensionality reduction* techniques, one is effectively isolating the differences between a stylistic and a neutral movement. Thus, by inverting the computational machinery used to recognize the movement one can obtain the ability to synthesize new movements which contain the features compatible with the desired style.

Computational and Biologically Inspired Models of Action Perception and Recognition

Early models of biological motion processing were closely tied to the point-light displays of Johansson. These models took as their input the image coordinates of the body points in successive frames and attempted to solve a series of equations for the three-dimensional structure of the point lights. The operation of these algorithms was essentially to incorporate the image coordinates within constraints such as the planarity of groups of points, or the hierarchical structure of points. These structure-from-motion calculations were essentially data driven (i.e., not requiring any information about body structure except for that incorporated into the computational constraints) and provided a means to explain both perceptual organization as well as the perception of body structure. While later empirical work called into question these particular models [25], they are still appealing in their approach to simultaneously explain perceptual organization and recovery of body form.

The biologically inspired models have taken as a starting point that the human visual system appears to separate the processing of motion and static form early in the processing streams. Additionally, these largely independent streams appear to converge in a brain region, known as the posterior superior temporal sulcus (STS), that brain imaging studies [26] have shown to specialize in the processing of biological motion [27]. The modeling approaches have studied the instantiation of biologically plausible computations within a hierarchical processing framework of form and motion [28] or emphasized the potential for template matching mechanisms to organize point-light displays [29]. While these biologically inspired computational models are broadly consistent with human behavioral experiments they are exceedingly complex to test at the physiological level. However, current investigations of the responses of single cells are beginning to reveal how motion and form neurons

are organized in cortex and the form and motion image characteristics to which they respond [30].

The computational and biologically inspired models have focused on the early and mid levels of visual processing in the interpretation of biological motion. However, since obtaining a visual understanding of the actions of others has significant social significance there has been activity in trying to understand how deeper meanings such as goals and intentions of actions are recovered. While it is possible that this understanding arises simply from a visual matching process that involves increasingly elaborate representations of the visual signal, there is evidence that a direct-matching route works by directly mapping visual input into ones own behavioral repertoire of actions. These direct-matching models are largely inspired by the finding of brain networks that represent both the production and perception of goal directed actions [31]. Consistent with these models recent brain imaging experiments have found the functional representations of movement goals and movement kinematics to be differentially represented within these networks [32].

Computational models of biological motion have proven useful in many ways. Not only do they provide a compact means to express how recognition might occur but they often lead to testable hypotheses that can be explored with further experiments. They also, importantly, allow a common framework for describing biological motion perception that can span related efforts in neuroscience and experimental psychology to understand how actions are recognized.

Related Entries

- ▶ Evaluation of Gait Recognition
- ▶ Gait Recognition, Model-Based
- ▶ Human Detection and Tracking
- ▶ Gait Recognition, Motion Analysis for
- ▶ Surveillance

References

1. Johansson, G.: Visual perception of biological motion and a model for its analysis. Percept. Psychophys. **14**(2)(Oct), 201–211 (1973)
2. Blake, R., Shiffrar, M.: Perception of human motion. Annu. Rev. Psychol. **58**, 47–73 (2007)
3. Ullman, S.: The Interpretation of Visual Motion. MIT Press, Cambridge, MA (1979)
4. Hiris, E., Humphrey, D., Stout, A.: Temporal properties in masking biological motion. Percept. Psychophys. **67**(3), 435–443 (2005)
5. Lu, H., Liu, Z.: Computing dynamic classification images from correlation maps. J. Vision **6**(4), 475–483 (2006)
6. Thornton, I.M., Pinto, J., Shiffrar, M.: The visual perception of human locomotion. Cognit. Neuropsychol. **15**(6–8), 535–552 (1998)
7. Mather, G., Radford, K., West, S.: Low-level visual processing of biological motion. Proc. R. Soc. London, Ser. B – Biol. Sci. **249** (1325), 149–155 (1992)
8. Neri, P., Morrone, M.C., Burr, D.C.: Seeing biological motion. Nature **395**(6705), 894–896 (1998)
9. Pavlova, M., Sokolov, A.: Prior knowledge about display inversion in biological motion perception. Perception **32**(8), 937–946 (2003)
10. Shipley, T.F.: The effect of object and event orientation on perception of biological motion. Psychol. Sci. **14**(4), 377–380 (2003)
11. Shipley, T., Zacks, J. (eds.): Understanding Events: How Humans See, Represent, and Act on Events. Oxford University Press, Oxford (2008)
12. Stevenage, S., Nixon, M., Vince, K.: Visual analysis of gait as a cue to identity. Appl. Cognit. Psychol. **13**, 469–474 (1999)
13. Loula, F., Prasad, S., Harber, K., Shiffrar, M.: Recognizing people from their movement. J. Exp. Psychol.: Human Percept. Perform. **31**, 210–220 (2005)
14. Troje, N.F., Westhoff, C., Lavrov, M.: Person identification from biological motion: Effects of structural and kinematic cues. Percept. Psychophys. **67**, 667–675 (2005)
15. Cutting, J., Kozlowski, L.: Recognizing friends by their walk: Gait perception without familiarity cues. Bull. Psychonom. Soc. **9**, 353–356 (1977)
16. Troje, N.F., Westhoff, C., Lavrov, M.: Person identification from biological motion: Effects of structural and kinematic cues. Percept. Psychophys. **67**(4), 667–675 (2005)
17. Gunns, R.E., Johnson, L., Hudson, S.M.: Victim selection and Kinematics: A point-light investigation of vulnerability to attack. Journal of Nonverbal Behavior **26**(3), 129–158
18. Abernethy, B., Gill, D.P., Parks, S.L., Packer, S.T.: Expertise and the perception of kinematic and situational probability information. Perception **30**(2), 233–252 (2001)
19. Johnson, K.L., Tassinary, L.G.: Perceiving sex directly and indirectly – meaning in motion and morphology. Psychol. Sci. **16** (11), 890–897 (2005)
20. Pollick, F., Lestou, V., Ryu, J., Cho, S.B.: Estimating the efficiency of recognizing gender and affect from biological motion. Vision Res. **42**(20), 2345–2355 (2002)
21. Cutting, J.E., Proffitt, D.R., Kozlowski, L.T.: A biomechanical invariant for gait perception. J. Exp. Psychol.: Human Percept. Perform. **4**(3), 357–372 (1978)
22. Pollick, F., Kay, J., Heim, K., Stringer, R.: Gender recognition from point-light walkers. J. Exp. Psychol.: Human Percept. Perform. **31**(6), 1247–1265 (2005)
23. Troje, N.F.: Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. J. Vision **2**(5), 371–387 (2002)

24. Brand, M., Hertzmann, A.: Style machines. In: SIGGRAPH 2000 Conference Proceedings, pp. 183–192. ACM, New York (2000)
25. Bertenthal, B.I., Pinto, J.: Global processing of biological motions. Psychol. Sci. **5**(4), 221–225 (1994)
26. Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., Blake, R.: Brain areas involved in perception of biological motion. J. Cognit. Neurosci. **12**(5), 711–720 (2000)
27. Oram, M.W., Perrett, D.I.: Responses of anterior superior temporal polysensory (stpa) neurons to biological motion stimuli. J. Cognit. Neurosci. **6**(2), 99–116 (1994)
28. Giese, M., Poggio, T.: Neural mechanisms for the recognition of biological movements.. Nat. Rev. Neurosci. **4**(3), 179–192 (2003)
29. Lange, J., Lappe, M.: A model of biological motion perception from configural form cues. J. Neurosci. **26**(11), 2894–2906 (2006)
30. Vangeneugden, J., Pollick, F.E., Vogels, R.: Functional differentiation of macaque visual temporal cortical neurons using a parametric action space. Cerebral Cortex Advance Access published on July 16, 2008
31. Rizzolatti, G., Craighero, L.: The mirror-neuron system. Annu. Rev. Neurosci. **27**, 169–192 (2004)
32. Lestou, V., Pollick, F.E., Kourtzi, Z.: Neural substrates for action understanding at different description levels in the human brain. J. Cognit. Neurosci. **20**(2), 324–341 (2008)

Punch-in, Clock-in, Punch-out, Clock-out, Punch

The term punch, describes the act of a mechanical strike putting a hole, or "punching" the timecard to signify the employees action of either registering for work, or leaving work. In an electronic world, this has changed to "clock-in" and "clock-out".

► Time and Attendance

Punctum Lacrimale

The punctum lacrimale is located at the corner where upper eyelid comes together with lower eyelid. It is prominent in many subjects and appears as a "D" shape. The function of punctum lacrimale is to secrete tears to keep proper moisture on the surface of eyeball.

► Iris Super-Resolution
► Automatic Classification of Left/Right Iris Images

Pupil

The pupil is a hole in the center of the iris that controls the amount of light entering the eye.

► Iris Image Data Interchange Formats, Standardization

Pupil Phase Engineered Iris Biometrics

► Wavefront Coded® Iris Biometric Systems

Pupil Phase Engineering

A general framework for the design of pupil phase masks for certain computational imaging systems, where high-quality image acquisition is addressed from an optimization perspective. Extending the depth of field is but one requirement of image quality, others being controlling and minimizing the impact of aberrations, motion blur, and scattering from the imaging medium, to name a few.

► Wavefront Coded® Iris Biometric Systems

Purkinje Images

When illuminating the eye the exterior and interior surfaces of the cornea and lens reflect the illuminating light forming bright reflections within recorded images. These are known as the Purkinje images. Within the field of biometrics Purkinje images are evident within iris images and can obscure areas of iris texture. Iris capture systems aim to minimize their effect reducing them in size and restricting them to within the pupil region.

► Simultaneous Capture of Iris and Retina for Recognition

Q

Quadrant

In Cartesian coordinate system, the intersection of the two axis (x and y) creates four regions, called quadrants. Conventionally, quadrants are labeled counterclockwise starting from the upper right ("northeast") quadrant. In the first quadrant, both x and y coordinates are positive. In the second quadrant, x-coordinates are negative and y-coordinates are positive. In the third quadrant, both coordinates are negative and in the fourth quadrant, x-coordinates are positive and y-coordinates are negative.

► Iris Recognition, Overview

Quality-dependent Fusion

► Fusion, Quality-Based

Quantum Efficiency (QE)

For imaging sensors, the probability that a single photon impinging on a detector will be detected by the sensor. The exact definition can vary from vendor to vendor, e.g., some vendors report the QE for photons that hit a sensor pixel, ignoring the fill factor – the fraction of the sensor that is actually covered by pixels.

► Iris Device

© 2009 Springer Science+Business Media, LLC

R

Radiometric Calibration

A process for achieving a direct relation between the value at a pixel and the absolute amount of thermal emission from the corresponding physical scene element.

► Face Recognition, Thermal
► Image Formation

RAIC

► Iris Recognition at Airports and Border-Crossings

Random Forgery

In signature verification, random forgeries (also known as simple forgeries) represent the case where forgers claim to be another user but use their own signature.

► Signature Databases and Evaluation

Range Scans

Data that has the 3D depth information of every scanned point. A scan is the reading of the information for a typically prespecified region. A range image is a collection of pixel values with corresponding depth information. In most instances, the sensor used to obtain the range image is calibrated allowing us to give the distance measures in physical units such as meters.

► Face Recognition, Component-Based

RASTA-Filtering

RASTA-filtering was originally introduced in connection with perceptual linear prediction (PLP) [4] type of preprocessing; i.e., band-pass filtering in the log spectral domain. It aims to suppress slow channel variations assumed to be additive. This filtering principle has also been applied to cepstral feature based preprocessing [21] in both the log spectral and the cepstral domains. A general RASTA filter is defined by:

$$T(z) = \frac{k \sum_{n=0}^{N} \left(n - \frac{n-2}{2}\right) z^n}{1 - px^{-1}}, \qquad (1)$$

where, the numerator is a regression filter of odd order N and the denominator is a leaky integrator. A simple variant of RASTA-filtering is a sliddly window mean subtraction technique, which corresponds to a moving average filter. Filtering is normally performed in the cepstral domain (CMS). The mean corresponds to the long-term cepstrum and is normally computed on the speech part. A silence/speech detector is thus necessary.

► Session Effects on Speaker Modeling

Raw Finger Vein Image

Raw finger vein image is the original infrared finger vein image captured by a finger vein reader. It is

© 2009 Springer Science+Business Media, LLC

typically a greyscale image in which finger vein patterns are imaged as dark area.

► Finger Vein Biometric Algorithm

Read Noise

For image sensors, each time when a pixel is read out, there is noise associated with the act of reading the pixel. This noise is, to first order, independent of the value of the pixel.

► Biometric Sensor and Device, Overview
► Iris Device

Real-Time 3D Surface Digitization

Many 3D biometrics applications rely on laser scanners for surface digitization. Although these devices produce very accurate 3D coordinates, they have two main limitations. They are expensive thus prohibiting their use in a wide scale, and they require the user to stay still during the acquisition. The latter is because during acquisition, a laser beam sweeps the object surface, and thus movement may results to artifacts. 3D acquisition based on the stereoscopic principle is an alternative without the above limitations, since two plain cameras are used. However the accuracy of such devices is usually limited. A solution that is a compromise between the above is using sensors that are based on the structured light approach. The surface of the object is illuminated by a light pattern (e.g. by means of a slide or video projector) and the resulting image is captured by a common camera. By analyzing the deformation of this pattern on the surface 3D coordinates may be computed by means of triangulation. The accuracy of this approach has been shown to be adequate for biometric applications.

► Finger Geometry, 3D

Recognition at a Distance

Recognition at a distance is the process of establishing the identity of humans in a non-intrusive way from a distance, often without their knowledge. Typically, face and gait biometrics are used to perform recognition from small to mid range distance.

► Face Recognition, Video-based

Rectilinear

Moving in, consisting of, bounded by, or characterized by a straight line or lines. Rectilinear images are those in which intensity values are recorded as a matrix, with each element in the matrix corresponding to the intensity measured by a single picture element (pixel) in the image sensor.

► Iris Image Data Interchange Formats, Standardization
► Iris Standards Progression

Reference Set

Reference set is a set of biometric samples or extracted features from a biometric trait of a user, which are stored to perform matching. Stored samples are also known as templates. Generally, each input biometric sample will be compared to all the templates in the reference set in the matching phase. Reference-based systems are opposed to model-based systems, where instead of storing genuine samples, a classifier is trained (or a statistical model is estimated) from the training set.

► Performance Evaluation, Overview
► Signature Matching

Reflection

When light impinges on a object, part of the light is reflected, it bounces off the object; part of the light is transmitted, it passes through the object; and part of the light is absorbed, it is converted from light into some other form of energy in the object.

The reflected light can be divided into two components, specular and diffuse. A specular reflection is what you see in a mirror; a single ray of light is reflected back as a single ray of light in a new direction determined by the rule that the angle of reflection with respect to a normal to the surface equals the angle of incidence with respect to the same normal. Diffuse reflection is what you see from a matte surface; a single ray of light is reflected back over a broad range of angles. Most materials have specular and diffuse components to their reflectivity.

► Iris Device

Reflection-Based Touchless Finger Imaging (RTFI)

RTFI refers to touchless fingerprint sensor.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Region-of-Interest (ROI) Encoding

It is often desirable to allocate a coding budget non-uniformly to an image, especially when compression is required. This may take the form of space-variant resolution, so that only some favored region receives maximal resolution while other areas are encoded with coarser resolution; or it may take the form of completely masking out the irrelevant areas. The capability to specify a Region-of-Interest (ROI) is built into advanced coding protocols for image compression, especially the JPEG2000 protocol (ISO Standard 15444) which allows binary masking as well as code-block selection so that different resolution levels and different coefficient scaling can be applied to different tiles. The older JPEG protocol (ISO Standard 10918), in its Part 3 extension, also supports variable quantization for explicitly specifying different quality levels for different image regions.

In JPEG2000, the MAXSHIFT tool allows specification of an ROI of arbitrary shape. ROI methods have been developed for compact encoding of iris images so that nearly all of the available coding budget is allocated to the iris itself, and not wasted on costly irrelevant structures such as eyelashes. In these methods, the non-iris regions of an image (eyelids, eyebrows, eyelashes, skin and sclera) are all automatically detected and painted out with uniform gray values before JPEG or JPEG2000 compression, resulting in almost no coding coefficients being wasted on them and allowing image file size to be reduced to 2,000 bytes.

► Iris Recognition Performance Under Extreme Image Compression

Registered Traveler

Catherine J. Tilton
Daon, Reston, VA, USA

Synonyms

Known traveler; Trusted traveler

Definition

Registered Traveler (RT) programs are designed to expedite legitimate travelers through a border control or security screening process by conducting a ► security threat assessment to determine risk levels prior to acceptance into the program and passage through a designated travel lane. These programs are generally

voluntary and involve the collection of biographic and biometric data used in background checks. Often, they involve the issuance of a secure biometric credential that can be used as part of an identity verification process.

Registered Traveler or Trusted Traveler programs are airline passenger security assessment systems deployed in the USA and controlled by the transportation security administration (TSA) as public/private partnerships. Their purpose is to expedite the security screening procedures for departing air passengers who have been previously vetted by TSA and deemed, on the basis of background checks, to be minimal security risks. To use the expedited fast lanes at airports, travelers must prove their identity and qualification under the scheme by either fingerprint or iris recognition, matching their own biometric record that is securely encrypted on a smartcard issued by TSA and interoperable across dozens of US airports.

The largest such programme is called CLEAR operated by Verified ID, which has enrolled more than 175,000 fee-paying members.

Introduction

Today's traveling public wants it all – hassle-free, secure, and safe travel. However, security usually means additional inconvenience. As a result, user-friendliness is sometimes seen as a trade-off for enhanced security. In the wake of the events of 9/11, national security and transportation officials worldwide were faced with this challenge: How to heighten the security of transportation, travel, and border systems while minimizing delays, aggravations, and privacy intrusions?

One idea that has gained some momentum is registering travelers in advance such that those travelers can be expedited through one or more of the travel processes – usually security screening and/or border control points. The idea is that by prescreening travelers, along with a strong identity authentication method, they can be segregated into risk categories, allowing security officials to allocate greater attention and resources to "unknown" or high-risk travelers [1].

Goals of RT programs generally include the following [2]:

- Enhance security
- Facilitate legitimate travel and trade
- Protect personal privacy

Participation in registered traveler programs involves a registration process in which the applicant provides biographic information and enrolls his/her biometrics. This information is used to conduct risk assessments (or security threat assessments) that may include criminal background and watchlist checks. The cognizant agency then makes an adjudication decision as to whether the individual is eligible for the program.

Benefits of participation from a traveler's perspective include shorter lines, faster processing, and consistency of experience. Some programs also offer other commercial benefits (e.g., preferred parking, concession discounts). When used to expedite security screening, one of the main benefits expressed by participants is predictability of wait times, allowing them to better judge time allowances and thus to spend more time at home or work prior to travel. This appeals mostly to frequent business travelers who are willing to pay for such an advantage. "The Privium (See "Example Systems") experience suggests that travelers are a great deal more willing to submit to fingerprint and iris scans if they think it will save them time and effort [3]."

RT programs exist in the US, Europe, and other locations and are emerging elsewhere. The US program centers on the security screening function, whereas many other programs focus on the border control area (i.e., immigration stations at ports of entry). Nearly all such programs are implemented at airports; however, this is not universally true. It is seen by many as a method for "simplifying passenger travel". A few such programs are highlighted in the examples section.

The RT Process

The processes used within a registered traveler program are similar to those used in typical biometric-based ▶ credentialing systems. (An example of a credentialing system is the personal identity verification program in the US.) [4] At a high level, the process comprises three elements: (1) Registration, (2) Authentication (Use), and (3) Administration. Depending on system architecture (see next section), these can be further broken down into component operations.

Registration is the process of enrolling and vetting an applicant to determine and instantiate eligibility. This generally involves the following steps:

- *Preenrollment* (optional). This step involves the applicant providing biographical information (usually via some remote means, such as the

internet) in advance of in-person, full enrollment. This may also include the scheduling of the enrollment appointment, receiving instructions, and payment of any associated fees.
- *Biographic enrollment.* The applicant either provides the requested biographic data or verifies/corrects those supplied during preenrollment. Biographic information would normally include name, date/place of birth, address, employment, and relevant identification numbers.
- *Identity proofing.* This entails the applicant providing one or more identity documents (e.g., birth certificate, drivers license, passport, or government issued ID), which are then scanned and validated.
- *Biometric enrollment.* Usually ten fingerprints as well as a facial photograph are captured, as these are the norm for criminal background checking. Some programs capture additional or alternative biometrics such as iris images.
- *Identity investigation.* Using the information provided by the applicant, one or more of the following checks may be performed as part of the security threat assessment (STA):
  - Name-based background checks, which may include verification of the existence of the claimed identity
  - Biometric duplicate checks, to ensure that an individual is not enrolled into the system under more than one claimed identity
  - Criminal history checks
  - Watchlist checks, including but not limited to known or suspected terrorists
- *Adjudication.* Based on the results of the STA, the cognizant authority makes a decision as to applicant eligibility.
- *Credential issuance (optional) and notification.* Most RT programs issue to approved participants a physical credential, such as a smartcard (or register a previously issued credential, such as an e-Passport) that is used in subsequent processes; however, this is not strictly required as a central server-based architecture is also possible. When used, these credentials serve as the claim of identity when presented, but also may contain biometrics that can be used to verify that claim during an authentication operation (this is discussed further in the following section).

Authentication is the process of verifying identity and program participation at the designated point in the travel process. This usually involves a claim of identity (either explicit or implicit), biometrically verifying that identity against the enrollment and validating the privilege of the participant. The last step may include:

- Checking the validity of the credential including expiration, revocation, integrity, and security features
- Checking that the participant's privileges have not been revoked or otherwise invalidated.

Authentication is frequently performed at (attended) verification kiosks located at RT privilege lines. The outbound users may need to first show a boarding pass and an ID (an RT card or a government-issued ID) before approaching the kiosk. He or she would then present his or her RT card e-Passport to the kiosk (in a credential based system), present his or her biometrics to the sensor when challenged, and be verified to pass through if his or her biometrics match. In most schemes, if biometric verification fails, the participant is moved the head of the regular queue. Some multibiometric RT systems allow the traveler to select their preferred verification biometric (e.g., fingerprint or iris).

Administration includes both normal system operation and maintenance functions as well as ongoing privilege management such as suspension or revocation of privileges, renewal and re-issuance, and sometimes periodic refresh of eligibility checks (i.e., running a participant back against the current watchlist).

Architecture

As you might expect, several architectural alternatives exist for RT schemes. Major architectural decisions include whether a centralized or distributed model will be used (this primarily relates to where the biometric matching operations are performed) and whether or not it will be federated (i.e., the degree of autonomy and commonality of design among service locations). Other decisions involve whether or not it will be credential based, single or multimodal, and whether a one-to-one (1:1) or one-to-many (1:N) technology will be employed.

In a centralized model, most major operations are performed at a central server. Participant facing operations are still performed locally (for example, information collection and biometric captures). However, all data storage/management and biometric matching are

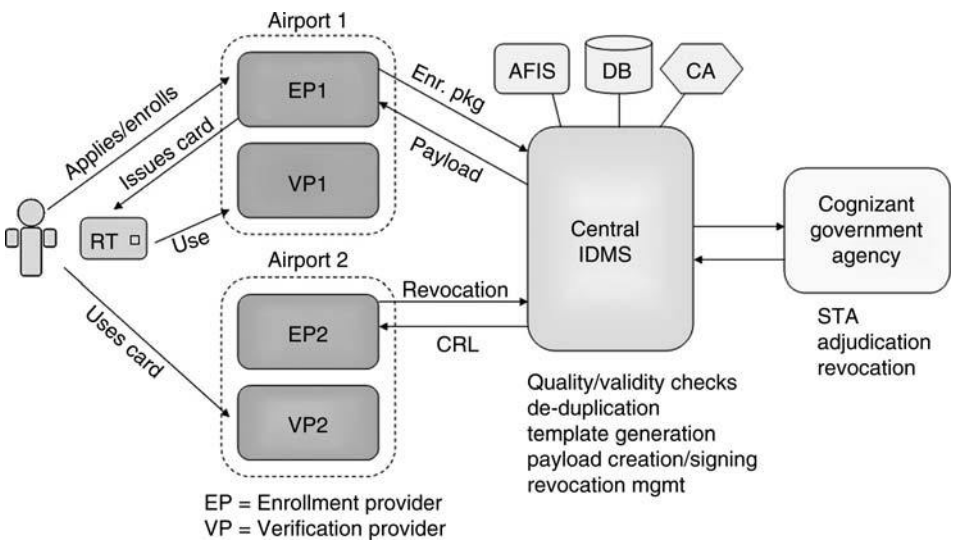
performed at the server. In a distributed model, data may be stored at the operational site(s) or on a credential and matching may be performed locally as well.

In a federated system, the architectural model is set and interfaces between major components are defined, but the local systems may be defined and controlled by multiple different entities, although some common requirements will apply (e.g., security). For example, in the US Registered Traveler system (see System Examples, below), which is setup as a public/private partnership, each enrollment and verification site is owned and operated by a given ▶ service provider (SP), which is contracted by a location sponsor (e.g., an airport). Design of the specifics of the enrollment station and verification kiosk, including capture equipment, process flow, and internal databases, is SP-specific, although elements that affect interoperability across the system are controlled by a system interoperability specification and conformance testing to that specification is performed. A central element still exists, of course, to perform those functions that require it (e.g., duplicate checking, STA gateway, payload generation/signing, etc.)

If a credential-based system is used, then there are a number of design elements involved including the card-operating system, data model, and card edge specification as well as whether the card personalization and issuance will be performed centrally, by an SP, or locally. Depending on the degree of federation, elements of the card topology (what it looks like on the outside) may also be strictly specified (e.g., whether/where a name and facial photo will be printed on the card). A very basic and important decision is whether a contact, contactless, or dual-interface card will be used. Security and privacy models must also be determined including how the personal information (including biometric data) will be protected, security features (topographical and digital), and tamper resistance. For contactless cards, this includes the prevention of covert reads and sniffing of RF communications during use. Figure 1 depicts a high level architecture and process flow for a federated, credential-based RT system.

In selecting the biometrics to be enrolled, stored, and verified, needs for background checking and operational use must be considered. Generally, background checking requires the collection of ten fingerprints (tenprints) as this is the standard for such systems; however, this may or may not be the biometric of choice for operational use. Considerations such as social acceptability, accuracy, and ease of use will drive this decision. To broaden the user population and reduce false rejections, a multimodal system may be employed. In this case, multiple biometrics are captured during enrollment and the participant may choose a preference at the time of enrollment or at verification. The need to be able to verify any of the set of supported modalities will, however, drive up the cost of the enrollment and verification stations as multiple capture devices will be required.

It is important to note that in order to perform duplicate (uniqueness) checking, the biometric collected



Registered Traveler. Figure 1 Federated RT Model.

for this purpose must be captured (or attempted to be captured) from all participants – otherwise, the purpose of such checking is defeated. Another important consideration for credential-based schemes is which biometrics will be stored on the credential itself, especially considering space limitations and data transfer rates of the card. This may constrain the number, types, and formats of biometric data to be held. For example, some image data (even compressed) may be too large to be feasibly considered, or if used, may further limit the inclusion of other biometrics. If fingerprint data are to be included, the number and selection of fingers must be determined.

Most registered traveler systems use 1:1 biometric verification; however, it is possible to use a 1:N identification technology (e.g., iris recognition). In this case, no claim of identity is made. The participant merely presents his or her iris(es) at the kiosk and this is matched against all other participants' records, resulting in an identification (identity returned as a result of a match) or a no-match. This is usually performed using a central server model to avoid the need to duplicate the database and protect all copies.

Today's systems are generally built around service-oriented architectures (SOA) taking advantage of the internet, existing SOA tools, and the many benefits of this architecture. These include service requester/provider decoupling, modularity, scalability, and component reuse. It is the basis of most large enterprise architectures and supports business process orchestration (BPO) workflow implementation, usually as part of an enterprise service bus (ESB).

Interoperability

In a nonfederated, single operator, closed RT system, interoperability is not generally an issue. However, in a federated system, this is critical as the various system components must be able to properly function and interact with one another. Areas in which interoperability are most critical include:

- Intrasystem (inter-subsystem) interfaces
- Card edge and data model
- Security

Examples of intersubsystem interfaces include those between an enrollment station/system and the central Identity Management System (IDMS), between the IDMS and the STA/adjudication agency, and between the IDMS and the verification station/system. The enrollment system must transmit collected "enrollment packages" to the IDMS for processing. The data and messaging formats as well as the communication protocols must therefore be defined. This includes the format of the biometric data. It is preferred that the submitted biometric information be in raw image format with minimal compression; however, some preprocessing and compression are usually required. For example, fingerprint data are normally transferred as three "slap" images in ANSI/NIST ITL1–2000/7 Type-14 record format using WSQ compression [5]. Messaging protocols are often Web services based, using a "SOAP over HTTP" XML-based implementation. In addition to the basic set of messages required to perform an enrollment, additional messaging is required to handle a host of error conditions and administrative needs. These include those related to updates, fees, and revocation.

With respect to an interoperable credential, the challenge in a federated system is that a credential produced by one service provider can be reliably and securely read and verified by a different provider. This generally requires a common form factor (e.g., ISO card), a common "card edge" or card interface/command set, and a common data model. The card application must be accessible and the security mechanisms usable by all authorized entities. A good model for this is the US PIV program already mentioned and the associated technical guidance, most notably NIST SP800–73 [6]. As an alternative, systems can leverage the e-Passport as an interoperable, biometrically-enabled credential, without the lost and logistics of issuing another. Common biometric data formats are also critical for interoperable use across multiple airport kiosks. For example, fingerprint data may be stored as ISO/IEC 19794–2 minutiae templates [7]. See also the chapter on Standardization for more information on biometric data interchange formats and technical interfaces.

Security is important in all RT systems, but becomes a bit trickier in federated systems due to the key management challenges. The intersubsystem messages should be signed and either encrypted or passed via an encrypted channel (i.e., SSL/TLS) using standard cryptographic protocols. Biometrics on a credential can be protected by one of the following means: PIN protection, biometric data encryption, and (card/reader) mutual authentication. In all the cases, the biometrics should be digitally signed either directly or

via a container hashing/signing scheme such as that used by ICAO for the ePassports.

Example Systems

This section takes a look at a few example registered traveler system implementations.

Privium

An early, and successful showcase, registered traveler program is the Privium system that was first introduced in 2002 at Schiphol airport in Amsterdam (following a one-year pilot), where participants are charged €99 for the privilege of bypassing the long queues at immigration (passport control).

This is a credential-based system where the iris data are stored on a smartcard issued to the participant. The enrolled iris data are not stored in a central database – only on the card, which also contains passport data. Any temporary use of this data is immediately purged after use to comply with Dutch privacy laws. Enrollment includes a background check. At the Privium verification station, a 1:1 biometric match is performed against the iris template stored on the Privium card.

The program is available to anyone with a European Economic Area (EEA) passport. It is said to cut the time spent at passport control to 10–15 s and queue time by up to 30 min [3]. Besides fast-track border passages, a separate check-in zone and priority parking are offered (Privium Plus). Participants may check in at the business class desk regardless of the class of ticket they hold. Online application and appointments are available. Onsite registration takes 15–20 min and includes inspection of identity documents, iris enrollment, and instructions on system operation. The Privium card is good for 4 years, but is renewed annually [8].

UK IRIS

The UK Home Office Border & Immigration Agency introduced its Iris Recognition Immigration System (IRIS) in March 2006, as a free service to the traveling public both to enhance their experience and to reduce manned immigration platforms. It is the first phase of the UK e-Borders' initiative to modernize immigration controls. Participation allows registered passengers to enter the UK without queuing to see an immigration officer at passport control [9]. The intent is to enhance both security and efficiency.

The Home Office says "Successfully enrolled passengers can enter the United Kingdom through automated immigration control barriers after looking into an iris recognition camera." Like the Privium program, IRIS uses iris recognition technology "because it is a fast, secure, and fraud-resistant way to verify passengers' identities. This makes it an ideal biometric for secure yet expedited clearance [9]". It is operational at Heathrow, Gatwick, Manchester, and Birmingham airports where enrollment rooms are provided. Enrollment takes approximately five to ten minutes.

Installation followed a trial in 2002 at Heathrow in which 200 selected passengers were enrolled. This was a joint project by British Airports Authority, the Immigration Service, British Airways, and Virgin Atlantic Airways. Figure 2 shows an IRIS verification station (barrier) at a UK arrival hall. Note the inclusion of multiple iris cameras mounted at various heights.

No fingerprint data are collected during IRIS enrollment – just iris and facial images – although an interview by an immigration official is conducted. No credentials are issued. Participation is voluntary and a marketing campaign is in place, targeting non-EEA foreign nationals. This has been a very popular system to date. As of February 2007, 61,000 people had registered. The original estimate when the program was announced was that over 1 million would be registered within the first five years.

US Registered Traveler

The United States piloted a registered traveler program in 2005 at five airports as part of a technology evaluation. Based on the results (including the popularity of the program which extended the original timeframe) and one additional private installation, a broader interoperable, national pilot system was deployed in late 2006. The US RT scheme is a fully fee-based public-private partnership. It uses a credential-based, federated architecture in which service providers contract with sponsors (airports/airlines) to operate the system. A central IDMS operated by the American Association of Airport Executives (AAAE) performs the functions identified in Fig. 1. The Transportation Security

Registered Traveler. Figure 2 UK IRIS Barrier.

Administration oversees the program and performs all eligibility adjudications.

As a federated system, an interoperability specification was developed by a group called the Registered Traveler Interoperability Consortium (RTIC). This group, which included airport operators, technology vendors, and system integrators with participation from the TSA, developed the specification over a 4 month period [10]. The TSA developed the program "to provide expedited security screening for passengers who volunteer to undergo a TSA-conducted security threat assessment (STA) in order to confirm that they do not pose or are not suspected of posing a threat to transportation or national security [2]." The process generally follows that described in the "Process" section earlier.

By December 2008, 20 airports had deployed the RT system and 200,000 participants had been enrolled, and over 2 million trips made through RT lanes. At this point, seven vendors had qualified as service providers. Participants can verify (against their RT card) using either fingerprint or iris recognition. Figure 3 depicts one of the US RT airport installations.

Canpass/Nexus-Air

In 2002, the Canadian Customs and Revenue Agency (CCRA) and Citizenship and Immigration Canada piloted an "express lane" through customs and immigration based on iris recognition. It is operational at airports in eight of the largest Canadian cities. The program allows preapproved low-risk air travelers to clear customs and immigration by using a self-service kiosk.

The goal was to streamline airport operations while maintaining a safe and secure border.

Now run by the Canada Border Services Agency (CBSA), this system was subsequently expanded via a bilateral Canada/US agreement into the NEXUS-AIR program which pre-clears travelers returning to the US from all major Canadian airports. CANPASS likewise pre-clears travelers returning from US (and some other international) locations.

Self-service kiosks use iris recognition to verify CANPASS membership. Through an agreement with CATSA, NEXUS-AIR members may also use the priority lane through the security screening checkpoint. In 2007, fees were $80 CDN. Key findings of a 2006 pilot evaluation found that participants were extremely satisfied with the program and reported saving an average of 27 minutes per passage. According to a survey conducted by EKOS Research Associates, the average NEXUS-Air participant enters Canada 12.8 times per year [11].

Others

RT systems have been implemented in other countries besides those highlighted earlier. For example, Germany has deployed a system in Frankfurt. Japan has recently

Registered Traveler. **Figure 3** US Registered Traveler System.

announced an RT component to their Japan Biometrics Identification (Border) system. And the EU has been investigating the possibility of an "international" RT program.

Issues and Considerations

Although apparently popular and workable, some issues and considerations exist with the establishment and use of RT systems.

Usability encompasses a range of issues related to ergonomics, accessibility, user interface, general ease of use, universality, and social acceptance. Biometric systems are typically not a "one-size fits all" arrangement, and care must be taken in their design to make them as usable as possible to the broadest possible user population. For example, nearly every biometric technology has a finite failure to enroll rate and some schemes are not able to be used by disabled individuals.

Privacy concerns also must be dealt with sensitively. Systems must take privacy considerations into account from the earliest point in system design and governance. Fortunately, this seems to be the case in the examples noted earlier.

Criticism has been levied against RT schemes as being a "pay to go to the head of the line" rather than a security improvement program – a so-called "Lexus line". "Not everyone is comfortable with what amounts to a two-tier system: some say it unfairly disadvantages infrequent travellers, and may make them susceptible to undue scrutiny [12]."

Summary

Biometric-based registered traveler programs are not only becoming more popular but also being proved to be workable and to provide benefits both to the implementers (i.e., governments and airports) and to the traveling public. Although similar in many areas, a variety of different schemes, processes, and architectures are employed to implement these programs. Biometrics are used both in the registration process, as part of the eligibility determination, and in the authentication process, during operational use to verify identity as a precleared program member.

Related Entries

- ▶ Border Management Applications
- ▶ Iris Recognition at Airports and Border-Crossings
- ▶ Standardization
- ▶ System Design

References

1. Registered Traveler Forum: What is an RT Program? (undated, online), http://www.registered-traveler-forum.com/
2. Transportation Security Administration: Registered Traveler page: http://www.tsa.gov/what_we_do/rt/index.shtm
3. Taylor, R.: All Eyes on Airport Security. The Guardian, 29 April 2004, http://www.guardian.co.uk/travel/2004/apr/29/travelnews1
4. Federal Information Processing Standard (FIPS) 201–1, Personal Identity Verification (PIV) of Federal Employees and Contractors, March 2006, NIST, http://csrc.nist.gov/groups/SNS/piv/index.html
5. ANSI/NIST ITL1–2000, Data Format for the Interchange of Fingerpint, Facial, & Scar Mark and Tatoo (SMT) Information, July 2000
6. NIST Special Publication 800–73–1, Interfaces for Personal Identity Verification, April 2006, http://csrc.nist.gov/publications/nistpubs/800–73–1/sp800–73–1v7-April20–2006.pdf
7. ISO/IEC 19794, Biometric Data Interchange Formats, Parts 1–10
8. Amsterdam Airport Schiphol, Privium Program website: http://www.schiphol.nl/privium/privium.jsp
9. Home Office, Border & Immigration Agency, IRIS project website: http://www.ind.homeoffice.gov.uk/applying/iris/
10. Registered Traveler Interoperability Consortium, RTIC Technical Interoperability Specification, Ver 1.2, 2 May 2007, http://www.rtconsortium.org/
11. Canada Border Security Agency (CBSA), NEXUS Air Pilot Project Evaluation Study, http://cbsa-asfc.gc.ca/agency-agence/reports-rapports/ae-ve/2006/nexus_air-eng.html#note_1
12. Calder, S.: How to bypass passport queues and leap the UK immigration hurdle in the blink of an eye. The Independent, 13 Aug 2007, http://travel.independent.co.uk/biztravel/article2872060.ece

Remote Authentication

Judith Markowitz
J. Markowitz, Consultants, Chicago, IL, USA

Synonyms

e-authentication, Remote access (partial); Remote monitoring (partial); Remote verification

Definition

As with any form of authentication, remote authentication involves the verification of a person's claim of identity. What makes remote authentication different is that it is performed on individuals located beyond the physical boundaries of the organization – what some call as the ▶ extended enterprise.

Remote authentication is distinct from contactless and at-a-distance. Those two methods refer to the absence of direct contact or proximity between the individual being authenticated and a biometric sensor. Their focus is on the physical relationship between the individual and the sensor.

In contrast, remote authentication says nothing about the proximity of the user to the input device. Both the user and the device are outside the perimeter of the enterprise and may involve contact, contactless, or at-a-distance input of biometric samples.

Remote authentication is also distinct from remote identification. "Authentication" is a synonym for verification which means the individual has made a claim of identity and the function of the biometric system is to determine whether the claim is valid.

Introduction

Historically the boundaries of enterprises, whether they were private corporations or government agencies, were largely defined by brick and mortar. Anyone wanting to do business with the enterprise would appear in person and their identity would be validated in situ. The primary forms of remote access were via wireline telephone and mail, and remote authentication was performed by humans.

Today, communication, transactions, and access to information are available to anyone who has any kind of telephone or Web-enabled computing device. This extended enterprise is an essential component of e-government and e-business which, in turn, are fueled by the Internet, globalization, and the mobile revolution. Furthermore, the individuals accessing the enterprise remotely include employees as well as partners, suppliers, customers, and the public. Furthermore, enterprises must specify the policies and procedures that govern the access granted to each of these individuals and groups.

Access to the extended enterprise has become increasingly electronic and dominated by automation, including ▶ interactive voice response (IVR) systems, email, chat, SMS, and Web sites.

The growth of remote access has exposed the failure of ID + password/PIN security to prevent or even attenuate unauthorized access to personal data. There has been a global escalation of identity theft and the perpetration of new types of fraud [1]. In response, security-enhancing regulations have been promulgated

at national and international levels. In the United States, alone, regulators now mandate multi-factor security for financial services [2–4], healthcare [5], and telecommunications [6].

The rise of remote access, concerns about the ability of existing authentication methods to prevent unauthorized access, regulation, and other factors have enhanced the attractiveness of biometrics for automating remote authentication. According to the National Institute of Standards and Technology's (NIST) publication *Introduction to Public Key Technology and the Federal PKI Infrastructure* [7] "biometrics provide a very high level of security because the authentication is directly related to a unique physical characteristic of the user which is more difficult to counterfeit." (NIST SP 800–32 pp. 8–9.) Biometrics is ideally suited to remote authentication because it binds the authentication event to the physical presence of the human claimant – even when the event is performed with a remote user.

Levels of Assurance

Any application requiring authentication can be described in terms of its security level, the degree of potential harm or impact that an authentication error would cause (from minor inconvenience to criminal offenses and threat to personal safety), and the likelihood that such harm or impact will occur. Those considerations become extremely important when the application must authenticate remote users. In response to concern about such issues the United States Office of Management and Budget (OMB) *Publisher E-Authentication Guidance for Federal Agencies* [8] to assist U.S. federal agencies implementing the "e-Gov" program.

OMB's Guidance defines four levels of authentication security (called "assurance levels"). Table 1 correlates each of the levels with the damage that could occur as the result of an authentication error and the potential harm to the enterprise.

The National Institute of Standards and Technology (NIST) translated the OMB's assurance levels into electronic authentication solutions and published them as *Electronic Authentication Guideline* (NIST SP 800–63) [9]. Technical committee MI (Biometrics) The American National Standards Institute/InterNational Committee for Information Technology Standards (ANSI/INCITS) extended NIST's categories to include biometrics and published its work as *Study Report on Biometrics in E-Authentication* (M1/07–0185) [10].

The following provides a summary of assurance levels in those three publications. The italicized description of each level is from Appendix A of OMB 04-04 [8]

- *Level 1: Little or no confidence in the asserted identity's validity* is required. (OMB Appendix A p. 3.) There needs to be some assurance that the claimant is the same person who originally registered. A single authentication method is sufficient, such as a plaintext password. No cryptographic methods are required and no effort is invoked to prevent an eavesdropper from discovering a password or other secret that might be used for authentication.
  Biometrics: "It is likely biometric technologies used alone would be stronger than the necessary security at this level." (American National Standards Institute/InterNational Committee for Information Technology Standards (ANSI/INCITS) 2007 *Study Report on Biometrics in E-Authentication* (M1/07–0185). p. 16).

Remote Authentication. Table 1 Maximum Potential Impacts for Each Assurance Level (from OMB M 04-04) [2]

Potential Impact Categories for Authentication Errors	Assurance Level Impact Profiles			
	1	2	3	4
Inconvenience, distress or damage to standing or reputation	Low	Mod	Mod	High
Financial loss or agency liability	Low	Mod	Mod	High
Harm to agency programs or public interests	N/A	Low	Mod	High
Unauthorized release of sensitive information	N/A	Low	Mod	High
Personal safety	N/A	N/A	Low	Mod High
Civil or criminal violations	N/A	Low	Mod	High

Examples:

- A user presents a self-registered name and user ID (e.g., a password) as part of creating a personalized space on a Web site or in a social network (e.g., MyYahoo or Facebook) where little or no sensitive, personal information is stored.
- A user participates in an online discussion group that does not request identifying information beyond participant name and organization name.

- *Level 2: Some confidence in the asserted identity's validity* [2] is required. This level requires only single-factor authentication plus approved cryptographic techniques. These credentials are appropriate for operations that require, a user's identity details be verified independently during the initial registration (Bolton, Joshua 2003 E-Authentication Guidance for Federal Agencies (Memorandum M-04-04) Office of Management and Budget (OMB) Appendix A p.3).
  Biometrics: "There is a contention that biometrics cannot be considered secrets and therefore there is language in this assurance level that prohibits the sharing of secrets. This limitation can be overcome, however, if there are countermeasures put in place to mitigate the concerns about the sharing of authentication secrets. In particular, through liveness detection at the point of acquisition and the use of approved cryptographic techniques to protect transmission." (ANSI/INCITS Op. cit. p. 16.) A biometric would be suitable and could be stronger than the security required for this level, especially content-bearing biometrics, such as ▶ text-dependent speaker authentication because they are two-factor techniques (biometric + knowledge).
  Examples:
  - A user subscribes to an online educational service that must authenticate the person in order to present the appropriate course material, assign grades, or demonstrate that the user has satisfactorily completed the training. The primary risk is that an unauthorized third party may gain access to the grades.
  - A beneficiary changes her or his address of record through the web site of an insurance company. The primary risk is missing mails that are sent to the beneficiary's address. The insurance company must assess the risk that an unauthorized individual would access the information.
- *Level 3: High confidence in the asserted identity's validity* [2] is required. This level requires a minimum of two authentication factors (e.g., a one-time password and a biometric). The claimant must demonstrate that she or he controls the authentication devices (called "tokens"). Cryptography must be used to protect the authentication token against man-in-the-middle, replay, and other attacks.
  Biometrics: "Assurance Level 3. . . specifically calls out the use of biometrics as an option in order for the claimant to prove that he or she controls the token." (ANSI/INCITS Op. cit. p. 17).
  Examples:
  - A patent attorney electronically submits confidential patent information to the United States Patent and Trademark Office. Improper disclosure would give competitors a competitive advantage.
  - A bank customer uses online or telephone banking to access account information or transfer restricted amounts of funds.
  - A corrections agency must monitor non-violent criminal offenders in home-incarceration and community-release programs.
  - An employee or contractor uses a remote system giving her or him access to sensitive, personal client information. The transactions occur over the Internet. The sensitive personal information available to the employee creates a moderate potential impact for unauthorized release.
- *Level 4: Very high confidence in the asserted identity's validity* [2]. This level requires a minimum of two-factor authentication that employs the strongest authentication and cryptographic techniques that can be applied to remote access.
  Biometrics: "Assurance Level 4 still requires two-factor authentication and does not prohibit the use of biometrics as an option in order for the claimant to prove that he or she controls the token." (ANSI/INCITS Op. cit. p. 17).
  Examples:
  - A law enforcement official accesses a law enforcement database containing criminal records. Unauthorized access could raise privacy issues and/or compromise investigations.

- A bank customer uses online or telephone banking to transfer unrestricted amounts of funds
- A physician uses a remote system to access the medical records of a patient.

Architectures for Remote Authentication

Architectures used for biometric remote-authentication range from highly centralized to widely distributed. There are two defining elements of those architectures and they are:

- Where reference template/models are stored
- Where matching is performed.

ANSI/INCITS' *Study Report on Biometrics in E-Authentication* [10] examines four storage and matching locations:

- Central Server
- Client/Workstation
- Device/Sensor
- Physical Token

Central Server

It is a centrally-located computer, sometimes called the "biometric authentication server." In remote authentication architectures the server need not be co-resident with the other resources involved in the authentication operations, including the biometric verification engine.

Storage

The most widely-used storage mechanism for remote authentication is to house all biometric reference models in a centralized repository that supports authentication from multiple locations. This approach

1. Allows the enterprise to maintain control over the reference templates/models
2. Facilitates frequently recurring operations, notably
   Adding templates/models for new users
   Updating data in the existing templates/models (adaptation)
   Deleting templates/models for users who have been removed from the system
3. Helps ensure a consistent level of security for all templates/models in the repository

Conversely, use of a centralized repository places the responsibility for security and privacy on the enterprise. The enterprise must institute policies and procedures that ensure the integrity and validity of the data in the repository. If matching is not performed locally, there is a danger that transmissions may be intercepted and used for replay attacks. Although data management for a centralized repository is generally easier than management of distributed resources, centralized repositories can become extremely large and unwieldy.

Matching

The reference template/model and the live biometric sample obtained from the claimant are compared/matched with the server. That server may or may not also house the centralized repository of biometric templates/models.

Centralized matching is useful when biometric input is highly distributed and is a logical option when the reference templates/models are stored in a centralized repository. As with centralized storage, centralized matching assigns control and responsibility to the enterprise. In particular, transmission of biometric samples over networks makes them vulnerable to network-based attacks.

Client/Workstation

A general-purpose workstation, usually a desktop or laptop PC. PDAs and other mobile devices may be included even though such use blurs the boundary between "client/workstations" and "devices." The client is the location/resource where users initiate the remote authentication process but none of these clients is a dedicated biometric authentication server even if they house other resources used for authentication.

Storage

One ore more reference templates/models are stored in a local repository. That repository may or may not be the same as the client used to access the authentication service. This approach is useful when a small number of users can access specific data or transactions. It eliminates a central point of attack for intruders and reduces the problem of managing an unwieldy repository. Local storage also makes it possible for a system to operate even when the central network connection is unavailable.

On the other hand, there is less control over the integrity of the data on the local machine. Also, local storage of templates/models does not support authentication from multiple locations unless there are multiple copies of templates/models. Management of multiple copies adds complexity to the system and could lead to administration and synchronization problems. For example, it may not be possible for the administrator to determine whether there are multiple enrollments for a single individual under the same or different names. Removal and updating can be more difficult which could leave sensitive data and systems vulnerable when, for example, a disaffected employee's templates/models are not purged from the entire system.

Matching

The reference template/model and the live biometric sample obtained from the claimant are compared/matched on the local workstation. If the reference template/model is stored in a central repository it must be downloaded to the local machine. If storage and matching are both performed on the same machine there is less chance of network-based attacks on the biometric data but there is the possibility of having attacks on the reporting of authentication decisions. Those possibilities arise from potential malware on the workstation or an attack on the network.

Device/Sensor

It is a biometric input device. It may be part of a larger peripheral that is attached to a workstation; embedded in a workstation or other device, (e.g. a keyboard or a cell phone) or it may be a telephone. If it is a "dumb" device, it does nothing more than capture and transmit raw or slightly-processed biometric data (e.g., a standard telephone). If it has more intelligence, it may be able to store or process data or perform matching.

Storage

Storage on the device provides rapid access if it is a device the user controls (e.g., a cellphone Vs. a dedicated biometric sensor) it can give the user control over the biometric template/model. Depending on the device, it may be possible to incorporate cryptographic methods to further secure the stored template/model.

Template/model deletion can be challenging but could be accomplished by the system administrator as part of a provisioning process. Performing deletion in this way requires the devices to be centrally managed which is easier to accomplish when the input device is a dedicated biometric sensor unit rather than a cellphone.

Matching

Matching on a biometric sensor involves embedding biometric technology into the device in a way that allows it to operate as a stand-alone system. The matching process is generally fast because there may be little or no communication with the outside until the matching process is complete. As a result, network-based attacks are eliminated as a threat.

Physical Token

An object capable of storing a biometric reference template/model and possibly performing operations related to authentication, such as encryption, feature extraction, and even matching. Physical tokens generally have technology to resist tampering. Typical examples of physical tokens are smart cards, PCMCIA cards, USB memory sticks, and RF tokens.

Storage

Storage of a single user's reference template/model on a token gives the user maximum control and privacy protection. Only the user determines when her or his token is used and for what purpose. This form of storage also makes it possible for the token to be used at multiple locations.

Since the token and enrolled template/model are both controlled by the user, it is not possible to determine whether the user has enrolled multiple times under different identities. This not only makes it difficult for the enterprise to maintain a definitive list of users but also proves harder to prevent fraudulence.

Additional costs may come from the need to deploy the dual-purpose sensors. Each biometric sensor must not only be able to accept live samples but must also be capable of reading the stored template/model on the token. Another common cost factor is replacement of lost or damaged tokens.

Matching

When both storage and matching are on the token, the opportunity for network attack is greatly diminished.

There are some vulnerability in the communication link between the smart card/token and the reader. Tokens capable of matching are more complex and more expensive to replace.

Kinds of Applications

The most widely-deployed speaker-authentication applications requiring remote authentication involve

- Data security
- Transaction security
- Remote monitoring
- Access security

at levels of assurance that vary with the sensitivity of the data and other factors. All four types of applications of remote authentication exist in real-world deployments.

Most of the examples provided in this section come from speaker authentication since it is the modality that exhibits the broadest spectrum of remote-authentication applications today.

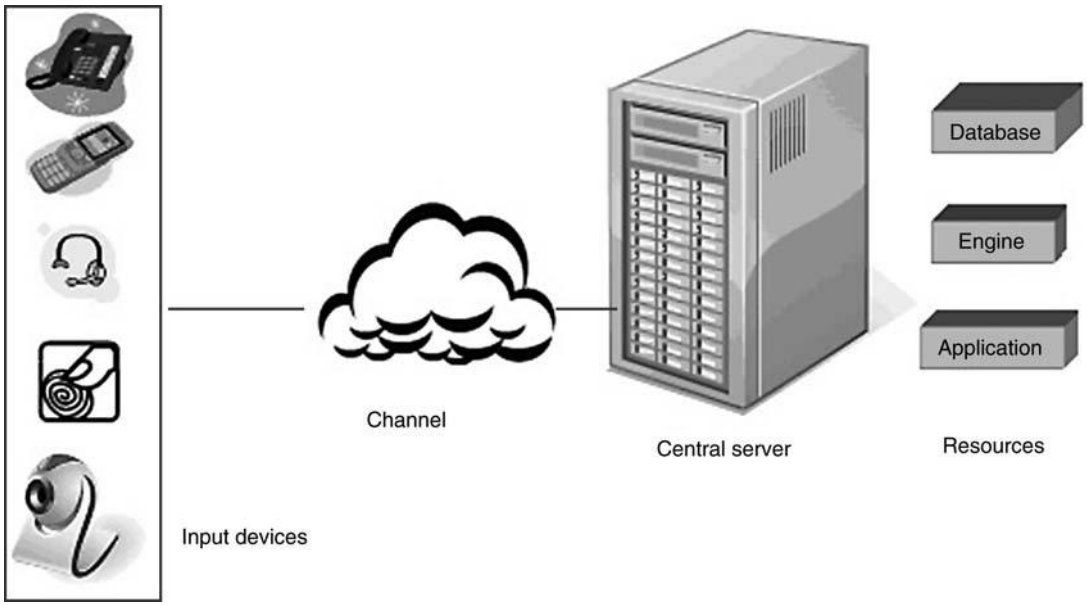
Speaker Authentication

Most remote applications of speaker authentication are centralized for both storage and matching (although the authentication resources including the application, biometric engine, and the template/model database) may not all reside on the same server. When authentication is hosted, for example, the application that calls the authentication process may reside within the enterprise but the biometric engine will be located on the server of the hosting-services provider. The template/models may be stored within the enterprise or by the hosting company.

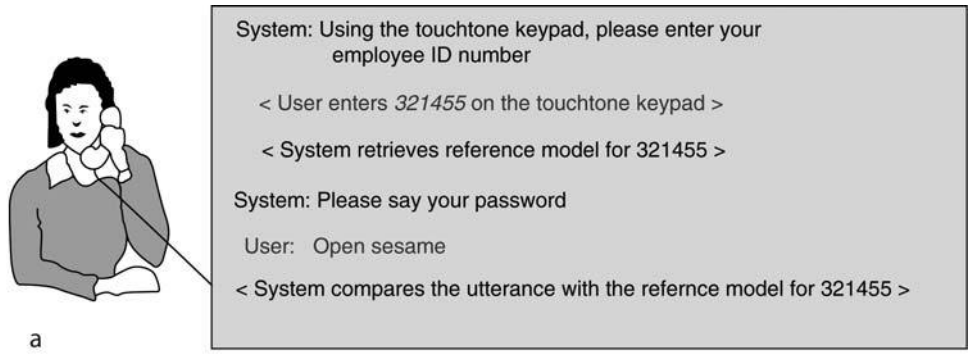
Figure 1 shows a typical centralized architecture. The remote-access channel, represented by the cloud in Fig. 1, can be a wireless or wireline telephone network, specialized data network (e.g., ATM), the Internet, or a combination of those channels. If the authentication is provided as a Web service, the cloud may include a Web browser. The assurance levels of these applications vary with the nature of the secured data, the resources involved, the users, and other factors.

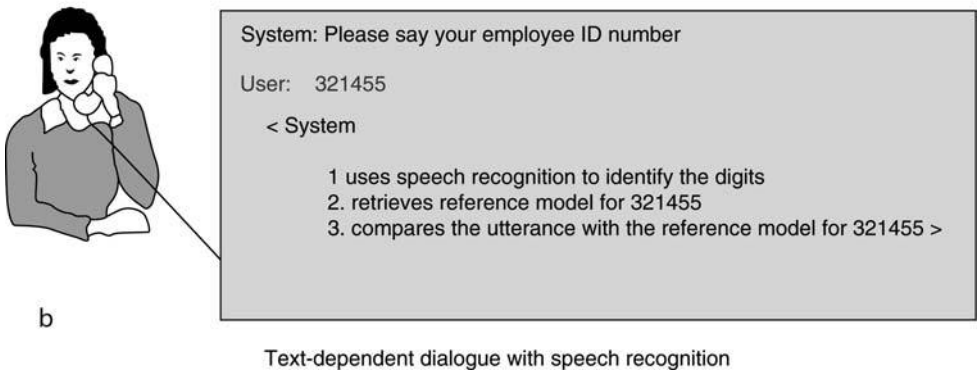
A typical deployment using speaker authentication is shown in Fig. 2. It employs two-factor authentication that combines password security with text-dependent, biometric security. The application consists of a dialogue involving interactive voice response (IVR). The IVR system answers the telephone and prompts the caller for an ID (claim of identity) and a spoken password (Fig. 2a). In some applications the system uses the spoken password as both the claim of identity and password (Fig. 2b). In those applications, ▶ speech recognition is used to decode the ID before it is sent to the authentication sub-system for biometric authentication.

The examples in Fig. 2 are typical for telephone banking (e.g., ABN AMRO, Banco Bradesco of Brazil,



Remote Authentication. Figure 1 Typical Centralized Architecture.





Remote Authentication. Figure 2 Examples of two-factor authentication using speaker authentication.

and Israel's Bank Leumi), account access (e.g., Bell Canada, Aeroplan, Australian Health Management, and Ameritrade), automated PIN/password reset (e.g., by Wells Fargo Bank, The Hartford Insurance, Swisscom, Morgan Stanley, VISA, AT&T, and Banco Santander International), and a range of other deployments, including

- Wells Fargo – credit card activation and customer helpdesk
- Bell Canada, CNRail, Telus – secure reporting, billing, dispatch instructions for field service personnel
- Austar (Australia) – allow club members to order movies
- United States Department of Homeland Security – telephone check-in and reporting by visa holders
- Municipality of Dubai – reports of littering offenses
- Union Pacific Railroad – customers report when their shipments are delivered (called "railcar release")
- Prisons – to ensure that inmates are not abusing their outbound-calling privileges.

These applications often include a "gray area" for matching-scores that fall slightly below the acceptance threshold. Scores that fall within the gray area trigger additional authentication procedures. Those procedures may include prompting for repetition of the password, a text-prompted (challenge-response) interaction, use of another authentication technology, or transfer to a human.

If a tape attack is suspected, the application may engage the user in a ▶ text-prompted interaction. Some deployments challenge the user to say something that she or he has never said to the system before, such as the answer to "What is today's date?"

Another approach using centralized architecture involves remote authentication of an individual calling the enterprise's call center and speaking with a human rather than interacting with an IVR. Most such deployments are designed to maintain a high level of authentication security while reducing the time needed to do the authentication. The agent initiates a ▶ text-independent session while speaking with a

caller when, for example, the caller requests a secured transaction, such as a sizable funds transfer, or sensitive account/customer information. While the system is running in the background, the agent may also be asking the caller questions (knowledge-based authentication). The combination of the factors produces the needed authentication. Bank Leumi, one of the largest banks in Israel, has used this approach for several years.

A method developed by Authentify, an American solutions provider, is suitable for authentication levels 3 and 4. It combines out-of-band voice authentication with a Web session. The following variant is used by VeriSign to authenticate applicants renewing digital certificates. The system sends an email to the applicant containing a link that initiates the authentication Web session. The system calls the telephone number provided by the applicant and performs speaker enrollment. The system then uses a third-party telephone directory to obtain the phone number for the applicant's company, calls that number, and asks to be connected with the applicant. When the applicant answers, the Web session displays a randomly-generated sequence on the applicant's computer screen and the telephone session asks the applicant to say that sequence and their name.

Electronic monitoring of community-released and home-incarcerated offenders utilizes text-prompted voice authentication. These systems place outbound calls to registered telephone numbers of locations where the offender is supposed to be (e.g., home, school, work, or AA meetings). The calls are placed at random times during the day and text-prompting is used to reduce the chance that offenders will use tape recorders.

Challenge response is also used for remote authentication employing other biometric modalities. This approach is used for employees, customers using networked devices (e.g., ATMs), and for registered airline travelers seeking to move quickly through security lines. Applications that require higher levels of authentication may request more than one fingerprint or samples for multiple biometric modalities.

Other Biometrics

A growing number of deployments are using smartcards with fingerprint, face, iris, or finger/hand vascular templates embedded in them. This includes e-passport, national ID, and trusted traveler programs. Matching is almost always done on the device or centrally (Fig. 2). One large-scale deployment in private industry is by ICICI Prudential Life Insurance of India. Its smartcard contains a fingerprint as well as the individual's policy information. Some biometric ATMs also use card-based storage.

Related Entries

- ▶ Biometrics, Overview
- ▶ Speaker
- ▶ Speaker Recognition, Standardization

References

1. Bosen, B.: Network Attacks: Analysis of Department of Justice Prosecutions 1999 – 2006. Pleasanton, CA: Trusted Strategies Ltd. (2006)
2. Authentication in an Internet Banking Environment (FIL 03–2005). United States Federal Deposit Insurance Corporation (2004)
3. Guidance on Authentication in an Internet Banking Environment. United States Federal Financial Institutions Examination Council (2005)
4. Q&A on Guidance on Authentication in an Internet Banking Environment. United States Federal Financial Institutions Examination Council (2006)
5. Health Insurance Portability and Accountability Act of 1996 (Public Law 104 191). United States Department of Health and Human Services (1996)
6. Telecommunications Carriers' Use of Customer Proprietary Network Information and Other Customer Information (FCC 07–22A). Federal Communications Commission (2007)
7. Kuhn, D.R., Hu, V.C., Polk, W.T., Chang, S.-J.: Introduction to Public Key Technology and the Federal PKI Infrastructure (SP 800–32). National Institute of Standards and Technology (2001)
8. Bolton, J.: E-Authentication Guidance for Federal Agencies. (Memorandum M-04–04, Appendix p. 3) Office of Management and Budget (OMB) (2003)
9. Burr, W.E., Dodson, D.F., Polk, W.T.: Electronic Authentication Guideline (SP 800–63 v 1.01.2). National Institute of Standards and Technology (2006)
10. Tilton, C., Young, M. (eds.): Study Report on Biometrics in E-Authentication (M1/07–0185). American National Standards Institute/International Committee for Information Technology Standards (2007)

Remote Monitoring (Partial)

▶ Remote Authentication

Remote Verification

▶ Remote Authentication

Rendering

▶ Face Sample Synthesis

Replay Attack

Replay attack is a term used in computer security where an attacker records a successful authentication procedure between a legitimate client and a computer system, or also between two computer systems, and then replays that recording in order to be falsely authenticated by the system. In the context of voice authentication, a replay attack involves a recording – analogue or digital – of a legitimate client's voice and the playing back of that recording to the authentication system by the attacker in order to be falsely accepted by the system as the legitimate client.

▶ Biometric Secutiry, Standardization
▶ Biometric Spoof Prevention
▶ Biometric System Design, Overview
▶ Liveness Assurance in Face Authentication
▶ Liveness Assurance in Voice Authentication
▶ Remote Authentication
▶ Security Issues, System Design
▶ Synthesis Attack
▶ Tamper-proof Operating System

Residence Time

The length of time a subject must reside in the capture volume of a biometric capture device to ensure that the device captures a good quality image.

▶ Iris Device
▶ Iris on the Move™

Resolution

In image analysis, a measure of the ability of a system to distinguish two features that are close together – to recognize that there are two features rather than one. In image displays, a measure of the ability of a system to present two features that are close together as distinct features rather than a single feature. In digital images and digital image processing, resolution is often described in terms of the number of pixels in the image or the number of pixels/unit length.

▶ Iris Device
▶ Photography for Face Image Data

Response Time

The time required by a biometric system to return a decision on identification or verification of a presented biometric sample. Response time includes the time for collecting data, extracting features, and matching against the enrolled biometric templates.

▶ Performance Evaluation, Overview

Retina

The retina is the multilayered sensory tissue of the posterior eyeball onto which light entering the eye is focused, forming a reversed and inverted image. It contains photosensitive receptor cells, the rods and cones, which are capable of converting light into nerve impulses that are conducted and further relayed to the brain via the optic nerve. There are about 110 to 125 million rods and 6.3 to 6.8 million cones in each human retina.

► Anatomy of Eyes
► Iris Image Data Interchange Formats, Standardization

Retina Recognition

Yoichi Seto
Advanced Institute of Industrial Technology, Tokyo Metropolitan University, Tokyo, Japan

Synonyms

Retinal scan; Vein Recognition; Vascular Recognition; Ocular biometrics

Definition

Retina recognition is a biometric technique that uses the unique patterns on a person's retina for person identification. The retina is the layer of blood vessels situated at the back of an eye. The eye is positioned in front of the system at a capture distance ranging from 8 cm to one meter. The person must look at a series of markers, viewed through the eyepiece, and line them up. The eye is optically focused for the scanner to capture the retina pattern. The retina is scanned with the near infrared (NIR 890 nm) irradiation and the unique pattern of the blood vessels is captured. Retina recognition makes use of the individuality of the patterns of the blood vessels. It has been developed commercially since the mid-1970s. Sandia Laboratory reported a false rejection rate of lower than 1.0%.

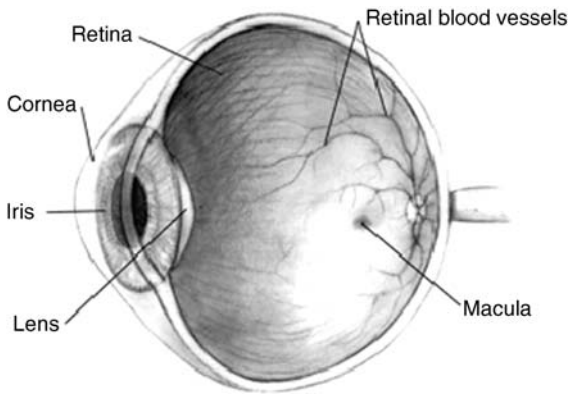
Introduction

The idea for retinal identification was first conceived by Dr. Carleton Simon and Dr. Isodore Goldstein and was published in the New York State Journal of Medicine in 1935, which while studying eye disease, made a study that every eye has its own totally unique pattern of blood vessels [1]. They subsequently published a paper on the use of retinal photographs for identifying people based on blood vessel patterns.

Referring to Fig. 1, the retina is to the eye as film is to the camera. Both detect incident light in the form of an image that is focused by a lens. The amount of light reaching the retina is a function of the iris. The retina is located on the back inside of the eyeball. Blood reaches the retina through vessels that come from the optic nerve. Just behind the retina there is a matting of vessels called the choroidal vasculature.

The retina is essentially transparent to the wavelength of light. The mat of vessels of the choroids just behind the retina reflects most of the useful information used to identify individuals [2, 3].

A retinal scan is used to map the unique patterns of a person's retina. The blood vessels within the retina absorb light more readily than the surrounding tissue and are easily identified with appropriate lighting. A retinal scan is performed by casting an undetectable ray of low-energy infrared light into a person's eye as they look through the scanner's eyepiece. This beam of light outlines a circular path on the retina. Because retinal blood vessels are more sensitive to light than the rest of the eye, the amount of reflection fluctuates.



Retina Recognition. Figure 1 Schematic representation of ball of the eye. Refer from Wikipedia.

The results of the scan are converted to a computer code and stored in a database.

Processing

The process of enrollment and verification/identification in a retinal scanning system is the same as the process for the other biometric technologies, which are (1) Acquisition and preprocessing of images, (2) Feature extraction, (3) Template data creation and (4) Matching processing [4].

Image Acquisition and Preprocessing

Illumination was provided by a narrowband near-infrared (NIR) LED. This wavelength selection was made for several reasons. Near-infrared light was not distracting to the subject and caused no visual discomfort. Additionally, it is important to choose wavelengths that give the best blood vessel contrast. The oxy-hemoglobin and reduction-hemoglobin can be found in the veins and arteries adequately absorb in the NIR. However, there are tradeoffs that one must consider when choosing NIR. A reduction-hemoglobin strongly absorbs the NIR irradiation energy, and then the part of vein pattern becomes dark.

The user must first place their eye onto a lens located in the retinal scanning device at an extremely close range. It is very important that the user must remain perfectly still at this point, in order to insure that a robust image will be captured. Also, the user must remove any eyeglasses that he or she might be wearing, because any light reflection from the lens of the eyeglasses could cause interface with the signal of the retinal scanning device. Once the user is situated comfortably, he or she then will notice a green light embedded against a white background through the lens of the scanning device. Once the retina scanning device is activated, this green light moves in a complete circle (360 degrees) and captures images of the blood vessel pattern of the retina through the pupil. At this phase, normally three to five images are captured. Also, this phase can take over 1 minute to complete, depending upon how cooperative the user is. This is considered to be a very long time in comparison to the image acquisition and processing times of the other biometric technologies.

The first encoding step was the identification of blood vessels within each image. Blood vessels were separated from distracters such as choroidal texture. The location and path of the retinal blood vessels were then quantitatively described. Sections of blood vessels were segmented and linked together. The identified blood vessel structure was then reduced to an efficient encoding template. Retina matching involves defining a similarity score between encoded blood vessel patterns. The encoding and final calculation of this similarity score must take into account the differences between the two source images.

Feature Extraction

A very strong advantage of retina recognition is that genetic factors do not dictate what the blood vessel pattern of the retina will be. This allows the retina to have very rich, unique features. As a result, it is possible that up to 400 unique data points can be obtained from the retina as opposed to other biometrics, such as fingerprint scanning, where only 30–40 data points (the minutiae) are available.

Template Data Creation

The unique features gathered from the blood vessel pattern of the retina forms the basis of the enrollment template, which is only 96 bytes, and as a result, is considered to be one of the smallest biometric templates.

Matching

The retinal matching involves defining a similarity score between encoded blood vessel patterns. The encoding and final calculation of this similarity score must take into account the differences between the two sources of images which is the same as the process for the other biometric technologies.

Sandia Laboratory has tested the retina recognition product of EyeDentify. The false rejection rate with databases of several hundred individual eyes is reported to be lower than 1.0% [3, 4].

Related Entries

► Back-of-hand Vein recognition
► Finger Vein recognition

► Palm Vein recognition
► Vein and Vascular recognition

References

1. Time.com: www.time.com/time/printout/0,8816,755453,00.html
2. Jain, A., Ruud, B., Sharath, P. et al.: Biometrics personal Identification in Network Society, pp. 123–141. Kluwer Academic Publishers, Dorrecht (1999)
3. Ruud, M.B., Jonathan, H.C., Sharath, P., Nalini, K.R., Andrew, W.S.: Guide to Biometrics, pp. 53–54. Springer, Heidelberg (2004)
4. Nanili, K.R., Venu, G.: Advances in Biometrics Sensor, Algorithms and Systems, pp. 133–155. Springer, Heidelberg (2008)

Retinal Angiogenesis

The formation of retina blood vessels by budding or sprouting from existing vessels. Random processes during retinal angiogenesis are thought to be responsible for the unique nature of the retinal blood vessel network.

► Simultaneous Capture of Iris and Retina for Recognition

Retinal Blood Vessels

The retina receives blood from two sources, the choroidal capillaries and the central retinal artery. The retinal and choroidal blood vessel pattern is unique to every person. The branches of the central artery and vein, for instance, diverge from the optic disc in a distinctive pattern that varies considerably across individuals. These retinal blood vessels are readily visible on a regular fundus photograph taken with visible light, whereas the choroidal blood vessels, forming a matting behind the retina, become visible when observed with near-infrared illumination.

► Anatomy of Eyes

Retinal Scan

► Retina Recognition

Reverse Engineering

Reverse Engineering refers mechanical disassembling and software analysis for architectural parsing for product to study/investigate operational mechanisms and its source code, etc.

► Embedded Systems

Revocable Biometrics

► Cancelable Biometrics

Ridge Enhancement

► Fingerprint Image Enhancement

Ridge Extraction

► Fingerprint Image Enhancement

Ridge Flow

The direction and overall pattern of a group of ridges in an area of friction ridge skin.

► Anatomy of Friction Ridge Skin

Robustness Test

Test to evaluate how much a certain influencing factor can affect biometric performance is robustness test.

► Influential Factors to Performance

ROC Curve

An ROC (receiver operating characteristic) curve is a plot commonly used in machine learning and data mining for exhibiting the performance of a classifier under different criteria. The *y*-axis is the true positive and the *x*-axis is the false positive (i.e., false alarm). A point on ROC curve shows that the trade-off between the achieved true positive detection rate and the accepted false positive rate.

► Face Detection
► Performance Measures

Rolled-Equivalent Fingerprint

It refers to a special impression of the fingerprint obtained by rolling the finger around the main finger axis on a planar surface.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Rolls Capture Device

It refers to a fingerprint device that allows the capture of rolled equivalent fingerprints. A special reconstruction algorithm is needed to compose the fingerprint during the rolling of the finger on the surface.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Rotation Angle

The rotation angle is the angle between the line joining the left and right pupil centers and the horizontal axis of the iris camera system. Counterclockwise rotation of the head about the optical axis of the camera is considered positive and clockwise rotation is considered negative.

► Iris Image Data Interchange Formats, Standardization
► Pose

S

Sample Quality

The intrinsic characteristic of a biometric signal may be used to determine its suitability for further processing by the biometric system or to assess its conformance to preestablished standards. The quality of a biometric signal is a numerical value (or a vector) that measures this intrinsic attribute (See also ► Biometric Sample Quality).

► Biometric Algorithms
► Fusion, Quality-Based

Sample Size

► Manifold Learning
► Performance Evaluation, Overview
► Test Sample and Size

Sampling Frequency

Sampling frequency is the number of samples captured in a second from the continuous hand-drawn signal to generate a discrete signal.

► Digitizing Tablet

Scalability

Scalability is the ability of a biometric system to extend adaptively to larger population without requiring major changes in its infrastructure.

► Performance Evaluation, Overview

Scenario Tests

Scenario tests are those in which biometric systems collect and process data from test subjects in a specified application. An essential characteristic of scenario testing is that the test subject is "in the loop," interacting with capture devices in a fashion representative of a target application. Scenario tests evaluate end-to-end systems, inclusive of capture device, quality validation software, enrollment software, and matching software.

► Performance Testing Methodology Standardization

Scene Marks

Crime scene marks are generally any physical phenomenon created or left behind and in relation to a crime scene, these can be fingerprints, blood spatter,

© 2009 Springer Science+Business Media, LLC

intentional and unintentional damage, or alteration to objects in the environment of the crime.

► Footwear Recognition

Scent Identification Line-Ups

Procedure where a trained dog matches a sample odor provided by a person to its counterpart in an array (or line-up) of odors from different people, following a fixed protocol. Scent identification line-ups are used in forensic investigations as a tool to match scent traces left by a perpetrator at a crime scene to the odor of a person suspected of that crime. The protocol includes certification of the team involved, collecting and conserving scent samples at crime scenes, collecting, conserving and presenting suspect, and other array odors, working procedures and reporting. Scent identification line-ups have evolved from simple line-ups that are used in human scent tracking/trailing, where a dog has to walk up to the person whose track it has been following and through some trained behavior indicate the person.

► Odor Biometrics

Score Fusion

► Fusion, Score-Level
► Multiple Experts

Score Fusion and Decision Fusion

Score fusion is a paradigm, which calculates similarity scores for each of the two modalities, then combines the two scores according to a fusion formula, e.g., the overall score is calculated as the mean of the two modality scores. Decision fusion is a paradigm, which makes an accept–reject decision for each of the two modalities, then combines the two decisions according to a fusion rule, e.g., the unknown sample is accepted only if both modalities yield an accept decision.

► Multibiometrics, Overview

Score Normalization

The score normalization techniques aim, generally, to reduce the scores variabilities in order to facilitate the estimation of a unique speaker-independent threshold during the decision step. Most of the current normalization techniques are based on the estimation of the impostors scores distribution where the mean, μ, and the standard deviation ν, depend on the considered speaker model and/or test utterance. These mean and standard deviation values will then be used to normalize any incoming score s using the normalization function

$$\text{score}N(s)\frac{s-\mu}{\nu}.$$

Two main score normalization techniques used in speaker recognition are:

1. *Znorm.* The zero normalization (Znorm) method (and its variants like Hnorm (Heck, L.P., Weintraub, M.: Handset-dependent background models for robust text-independent speaker recognition. In: ICASSP. (1997))) normalizes the score distribution using the claimed speaker statistics. In other words, the claimed speaker model is tested against a set of impostors, resulting in an impostor similarity score distribution which is then used to estimate the normalization parameters μ and ν. The main advantage of the Znorm is that the estimation of these parameters can be performed during the training step.
2. *Tnorm.* The test normalization (Tnorm) (Auckenthaler, R., Carey, M., Lloyd-Thomas, H.: Score normalization for text-independent speaker verification systems. Digital Signal Processing 10 (2000) 4254) is another score normalization technique in which the parameters μ and ν are estimated using the test utterance. Thus, during testing, a set of impostor models is used to calculate impostor scores for the given test utterance. μ and ν are estimated using these scores. The Tnorm is known to improve the performances particularly in the region of low false alarm.

Any of a number of rules for adjusting a raw similarity score in a way that takes into account factors such as the amount of data on which its calculation was based, or the quality of the data. One purpose of score normalization in biometrics is to prevent the arising of false matches simply because only a few elements (e.g., biometric features) were available for comparison. So an accidental match by chance would be more like tossing a coin only a few times to produce a perfect run of all head. Another purpose of score normalization is to make it possible to compare or to fuse different types of measurements, as in multibiometrics. For example, Z-score normalization redefines every observation in units of standard deviation from the mean, thereby allowing incommensurable scores (like height and weight) to become commensurable (e.g., he is 3.2 standard deviations heavier than normal but 2.3 standard deviations taller than normal). Frequently the goal of score normalization is to map samples from different distributions into normalized samples from a universal distribution. For example, in iris recognition a decision is made only after the similarity score (fractional Hamming Distance) has been converted into a normalized score that compensates for the number of bits that were available for comparison, thereby preventing accidental False Matches just because of a paucity of visible iris tissue.

► Score Normalization Rules in Iris Recognition
► Session Effects on Speaker Modeling
► Speaker Matching

Score Normalization Rules in Iris Recognition

John Daugman
Cambridge University, Cambridge, UK

Synonyms

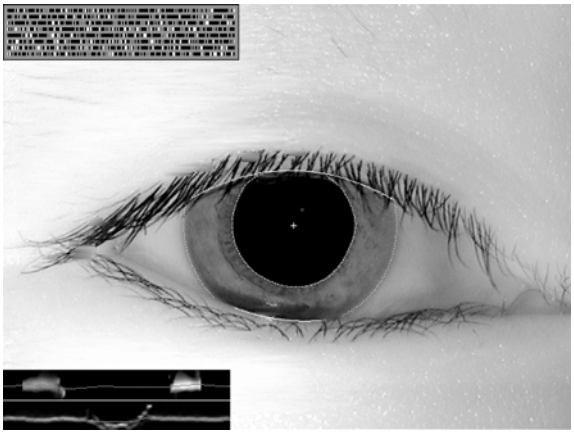
Commensurability; Decision criterion adjustment; Error probability non-accumulation; Normalised Hamming Distance

Definition

All biometric recognition systems are based on similarity metrics that enable decisions of "same" or "different" to be made. Such metrics require normalizations in order to make them commensurable across comparison cases that may differ greatly in the quantity of data available, or in the quality of the data. Is a "perfect match" based only on a small amount of data better or worse than a less perfect match based on more data? Another need for score normalization arises when interpreting the best match found after an exhaustive search, in terms of the size of the database searched. The likelihood of a good match arising just by chance between unrelated templates must increase with the size of the search database, simply because there are more opportunities. How should a given "best match" score be interpreted? Addressing these questions on a principled basis requires models of the underlying probability distributions that describe the likelihood of a given degree of similarity arising by chance from unrelated sources. Likewise, if comparisons are required over an increasing range of image orientations because of uncertainty about image tilt, the probability of a good similarity score arising just by chance from unrelated templates again grows automatically, because there are more opportunities. In all these respects, biometric similarity ► score normalization is needed, and it plays a critical role in the avoidance of False Matches in the publicly deployed algorithms for iris recognition.

Introduction

Biometric recognition of a person's identity requires converting the observed degree of similarity between presenting and previously enrolled features into a decision of "same" or "different." The previously enrolled features may not be merely a single feature set obtained from a single asserted identity, but may be a vast number of such feature sets belonging to an entire national population, when identification is performed by exhaustively searching a database for a sufficiently good match. The ► similarity metrics used for each comparison between samples might be simple correlation statistics, or vector projections, or listings of the features (like fingerprint minutiae coordinates and directions) that agreed and of those that disagreed as percentages of the total set of features extracted. For each pair of feature sets being compared, varying amounts of data may be available, and the sets might need to be compared under various transformations such as image rotations when the orientation is uncertain. An example is seen



Score Normalization Rules in Iris Recognition. Figure 1 Illustration of limited data being available in an iris image due to eyelid occlusion, as detected in a segmentation process.

in Figure 1, in which only 56% of the annular iris area is visible between the eyelids. Iris images may have also been acquired with a tilted camera (not unusual for handheld cameras), or with the head tilted or the eye rotated (cyclovergence) by an unknown degree, requiring comparisons to be made over a range of configurations for each of the possible identities, and with varying amounts of template data being available in each case. This article is concerned with the methods of ▶ score normalization that are used in iris recognition to make all of those comparison cases ▶ commensurable with each other, preventing False Match probability from rising simply because there is less data available for comparison or because there are many more candidates and match configurations to be considered.

Score Normalisation by the Amount of Iris Visible

The algorithms used in all current public deployments of iris recognition [2] work by a test of statistical independence: A match is declared when two templates *fail* the test of statistical independence; comparisons between different eyes are statistically guaranteed to pass that test [1]. The test of independence is based on measuring the fraction of bits that disagreed between two templates, called ▶ IrisCodes, and so the similarity metric is a ▶ Hamming Distance between 0 and 1. (The method by which an IrisCode is created is described in this encyclopedia in the entry on *Iris Encoding and Recognition using Gabor Wavelets.*) If two IrisCodes were derived from different eyes, about half of their bits should agree and half should disagree (since any given bit is equally likely to be 1 or 0), and so a Hamming Distance close to 0.5 is expected. If both IrisCodes were computed from the same eye, then a much larger proportion of the bits should agree since they are not independent, and so a Hamming Distance much closer to 0 is expected. But what is the effect of having varying numbers of bits available for comparison, for example, because of eyelid occlusion?

Eyelid boundaries are detected (as illustrated by the spline curve graphics in Figure 1 where each lid intersects the iris), and the parts of the IrisCode that are then unavailable are marked as such by setting masking bits. The box in the lower-left corner of Figure 1 shows Active Contours computed to describe the pupil boundary (lower "snake") and the iris outer boundary (upper snake). As these snakes are curvature maps, a circular boundary would be described by a snake that was flat and straight. The two thick grey regions in the box containing the upper snake represent the limited regions where the iris outer boundary is visible and possesses a large radial gradient (or derivative) in brightness. The gaps that separate the two thick grey regions correspond to parts of the trajectory around the iris where no such boundary is visible, because it is occluded by eyelids. Thus the outer boundary of the iris must be estimated (dotted curve) by two quite limited areas on the left and right sides of the iris where it is visible. In the coordinate system that results, the iris regions obscured by eyelids are marked as such by masking bits.

The logic for comparing two IrisCodes to generate a raw Hamming Distance HD_{raw} is given in Equation (1), where the data parts of the two IrisCodes are denoted $\{codeA, codeB\}$ and the vectors of corresponding masking bits are denoted $\{maskA, maskB\}$:

$$HD_{\text{raw}} = \frac{\|(codeA \otimes codeB) \bigcap maskA \bigcap maskB\|}{\|maskA \bigcap maskB\|} \quad (1)$$

The symbol $\otimes$ signifies the logical Exclusive-OR (XOR) operator which detects disagreement between bits; $\bigcap$ signifies logical AND whereby the masks discount data bits where occlusions occurred; and the norms $\| \ \|$ count the number of bits that are set in the result. Bits may be masked for several reasons other than eyelid or eyelash occlusion. They are also deemed

unreliable if specular reflections are detected in the part of the iris they encode, or if the signal-to-noise ratio there is poor, for example, if the local texture energy is so low that the computed wavelet coefficients fall into the lowest quartile of their distribution, or on the basis of low entropy (information density).

The number of bits pairings available for comparison between two IrisCodes, $\|maskA \bigcap maskB\|$, is usually almost a thousand. But if one of the irises has (say) almost complete occlusion of its upper half by a drooping upper eyelid, and if the other iris being compared with it has almost complete occlusion of its lower half, then the common area available for comparison may be almost nil. How can the test of statistical independence remain a valid and powerful basis for recognition when very few bits are actually being compared? It may well be that a less exact match on a larger quantity of data is *better* evidence of a match than is a perfect match on less data. An excellent analogy is a test of whether or not a coin is "fair" (*i.e.,* gives unbiased outcomes when tossed): Getting a result of 100% "heads" in few tosses (*e.g.,* 10 tosses) is actually much more consistent with it being a fair coin than getting a result of 60% / 40% after 1,000 tosses. (The latter result is 6.3 standard deviations away from expectation, whereas the former result is only 3.2 standard deviations away from expectation; so the 60/40 result is actually much stronger evidence against the hypothesis of a fair coin, than is the result of "all heads in 10 tosses".) Similarly, in biometric comparisons, getting perfect agreement between two samples that extracted only ten features may be much *weaker* evidence of a good match than a finding of 60% agreement among a much larger number of extracted features.

This is illustrated in Table 1 for an actual database of 632,500 IrisCodes computed from different eyes in a border-crossing application in the Middle East [3]. A database of this size allows 200 billion different pair comparisons to be made, yielding a distribution of 200 billion HD_{raw} similarity scores between different eyes. These HD_{raw} scores were broken down into seven categories by the number of bits mutually available for comparison (*i.e.,* unmasked) between each pair of IrisCodes; those bins constitute the columns of Table 1, ranging from 400 bits to 1,000 bits being compared. The rows in Table 1 each correspond to a particular decision threshold being applied; for example, the first row is the case that a match is declared if HD_{raw} is 0.260 or smaller. The cells in the Table give the observed False Match Rate in this database for each decision rule and for each range of numbers of bits being compared when computing HD_{raw}.

Using the findings in Table 1, it is informative to compare performance for two decision criteria: a very conservative criterion of $HD_{\text{raw}} = 0.260$ (the first row), and a more liberal criterion $HD_{\text{raw}} = 0.285$ (the sixth row) which allows more bits to disagree (28.5%) while still declaring a match. Now if the False Match Rates

Score Normalization Rules in Iris Recognition. **Table 1** False match rate without score normalisation: dependence on number of bits compared and criterion

HD_{Crit}	400 bits	500 bits	600 bits	700 bits	800 bits	900 bits	1,000 bits
0.260	$2 \cdot 10^{-9}$	$5 \cdot 10^{-10}$	$3 \cdot 10^{-10}$	$1 \cdot 10^{-10}$	0	0	0
0.265	$3 \cdot 10^{-9}$	$8 \cdot 10^{-10}$	$5 \cdot 10^{-10}$	$2 \cdot 10^{-10}$	$4 \cdot 10^{-11}$	0	0
0.270	$4 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$9 \cdot 10^{-10}$	$5 \cdot 10^{-10}$	$2 \cdot 10^{-10}$	0	0
0.275	$7 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$9 \cdot 10^{-10}$	$5 \cdot 10^{-10}$	$3 \cdot 10^{-11}$	0
0.280	$1 \cdot 10^{-8}$	$4 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$2 \cdot 10^{-10}$	0
0.285	$2 \cdot 10^{-8}$	$7 \cdot 10^{-9}$	$4 \cdot 10^{-9}$	$3 \cdot 10^{-9}$	$2 \cdot 10^{-9}$	$5 \cdot 10^{-10}$	$2 \cdot 10^{-11}$
0.290	$3 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$8 \cdot 10^{-9}$	$7 \cdot 10^{-9}$	$4 \cdot 10^{-9}$	$1 \cdot 10^{-9}$	$1 \cdot 10^{-10}$
0.295	$4 \cdot 10^{-8}$	$2 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$9 \cdot 10^{-9}$	$3 \cdot 10^{-9}$	$4 \cdot 10^{-10}$
0.300	$6 \cdot 10^{-8}$	$3 \cdot 10^{-8}$	$3 \cdot 10^{-8}$	$2 \cdot 10^{-8}$	$2 \cdot 10^{-8}$	$7 \cdot 10^{-9}$	$9 \cdot 10^{-10}$
0.305	$9 \cdot 10^{-8}$	$6 \cdot 10^{-8}$	$5 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	$1 \cdot 10^{-8}$	$2 \cdot 10^{-9}$
0.310	$1 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$8 \cdot 10^{-8}$	$8 \cdot 10^{-8}$	$7 \cdot 10^{-8}$	$3 \cdot 10^{-8}$	$5 \cdot 10^{-9}$
0.315	$2 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$6 \cdot 10^{-8}$	$1 \cdot 10^{-8}$
0.320	$3 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$1 \cdot 10^{-7}$	$2 \cdot 10^{-8}$

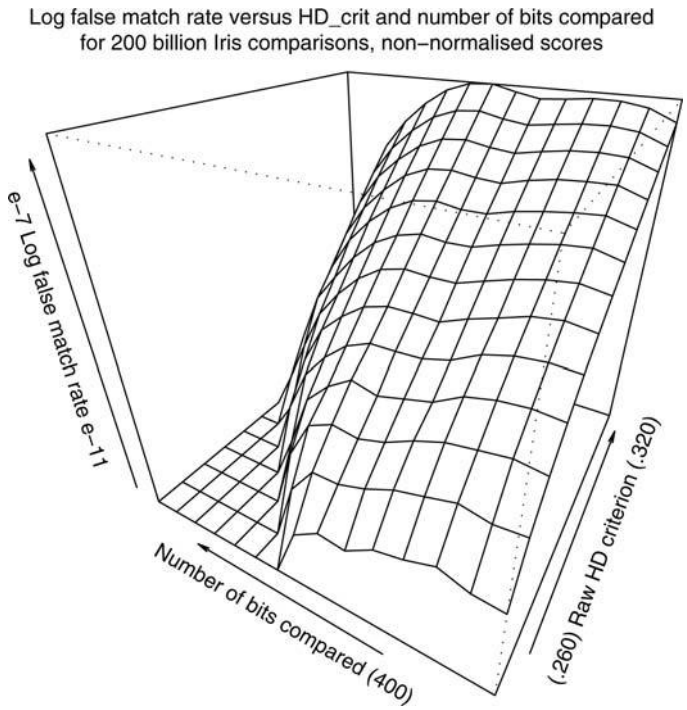
are compared in the first and last columns of these rows, namely when only about 400 bits are available for comparison and when about 1,000 bits are compared, it can be seen that, in fact, the more conservative criterion (0.260) actually produces 100 times more False Matches using 400 bits than does the more liberal (0.285) criterion when using 1,000 bits. Moreover, the row corresponding to the $HD_{\text{raw}} = 0.285$ decision criterion reveals that the False Match Rate is 1,000 times greater when only 400 bits are available for comparison than when 1,000 bits are compared.

The numerical data of Table 1 is plotted in Figure 2 as a surface, showing how the logarithm of the False Match Rate decays as a function of both variables. The surface plot reveals that there is a much more rapid attenuation of False Match Rate with increase in the number of bits available for comparison (lower-left axis), than by reduction of the HD_{raw} decision criterion in the range of 0.260 - 0.320 (lower-right axis). This is to be expected, given that iris recognition works by a test of statistical independence. The observations of Table 1 and Figure 2 clearly demonstrate the need for similarity scores to be normalized by the number of bits compared when calculating them.

A natural choice for the score normalization rule is to rescale all deviations from $HD_{\text{raw}} = 0.5$ in proportion to the square-root of the number of bits that were compared when obtaining that score. The reason for such a rule is that the expected standard deviation in the distribution of coin-tossing outcomes (expressed as a fraction of the n tosses having a given outcome), is $\sigma = \sqrt{pq/n}$ where p and q are the respective outcome probabilities (both nominally 0.5 in this case). Thus, decision confidence levels can be maintained irrespective of how many bits n were actually compared, by mapping each raw Hamming Distance HD_{raw} into a normalized score HD_{norm} using a re-scaling rule such as:

$$HD_{\text{norm}} = 0.5 - (0.5 - HD_{\text{raw}})\sqrt{\frac{n}{911}} \qquad (2)$$

This normalization should transform all samples of scores obtained when comparing different eyes into



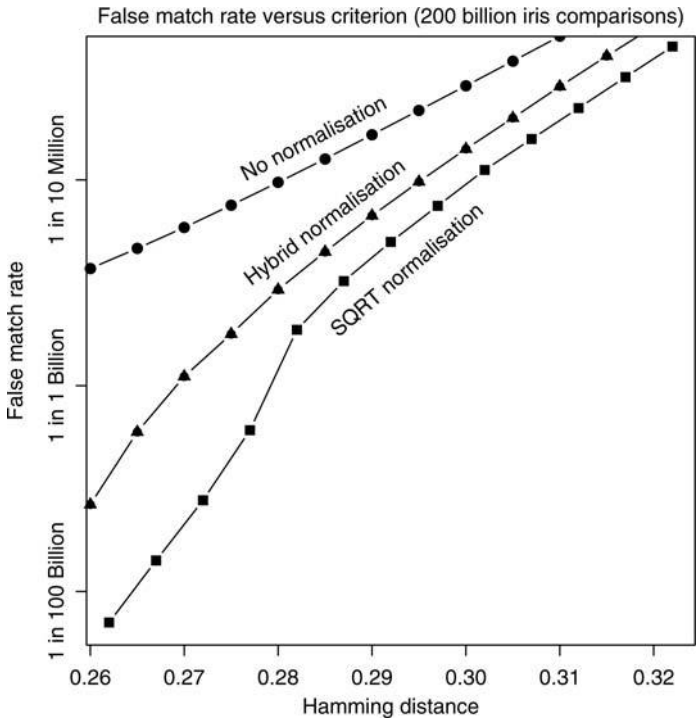
Score Normalization Rules in Iris Recognition. Figure 2 The data of Table 1 plotted as a surface in semilogarithmic coordinates, showing a range factor of 10,000-to-1 in the False Match Rate as the number of bits compared ranges from 400 to 1,000. This bit count is more influential than is the HD_{raw} decision criterion for unnormalised scores in the 0.260 - 0.320 range.

samples drawn from the same ▶ binomial distribution, whereas the raw scores HD_{raw} might be samples from many different binomial distributions having standard deviations σ dependent on the number of bits n that were actually available for comparison. This normalization maintains constant confidence levels for decisions using a given Hamming Distance threshold, regardless of the value of n. The scaling parameter 911 is the typical number of bits compared (unmasked) between two different irises.

The effect of using this normalization rule ("SQRT") is shown in Figure 3 for the 200 billion comparisons between different irises, plotting the observed False Match Rate as a function of the new HD_{norm} normalized decision criterion. Also shown for comparison is the unnormalized case (upper curve), and a "hybrid" normalization rule which is a linear combination of the other two, taking into account the number of bits compared only when in a certain range [4]. The benefit of score normalization is profound: it is noteworthy that in this semilogarithmic plot, the ordinate spans a factor of 300,000 to 1.

The price paid for achieving this profound benefit in robustness against False Matches is that the match criterion becomes more demanding when less of the iris is visible. Table 2 shows what fraction of bits HD_{raw} (column 3) is allowed to disagree while still accepting a match, as a function of the actual number of bits that were available for comparison (column 1) or the approximate percent of the iris that is visible (column 2). In every case shown in this Table, the probability of making a False Match is about 1 in a million; but it is clear that when only a very little part of two irises can be compared with each other, the degree of match required by the decision rule becomes much more demanding. Conversely, if more than 911 bits (the typical case, corresponding to about 79% of the iris being visible) are available for comparison, then the decision rule becomes more lenient in terms of the acceptable HD_{raw} while still maintaining the same net confidence level.

Finally, another cost of using this score normalization rule is apparent if one operates in a region of the ROC curve corresponding to a very nondemanding



Score Normalization Rules in Iris Recognition. **Figure 3** Comparing the effects of three score normalisation rules on False Match Rate as a function of Hamming Distance.

Score Normalization Rules in Iris Recognition. Table 2
Effect of score normalisation on the match quality required with various amounts of iris visibility

Number of bits compared	Approximate percent of iris visible (%)	Maximum acceptable fraction of bits disagreeing
200	17	0.13
300	26	0.19
400	35	0.23
500	43	0.26
600	52	0.28
700	61	0.30
800	69	0.31
911	79	0.32
1,000	87	0.33
1,152	100	0.34

False Match Rate, such as 0.001, which was the basis for NIST ICE (Iris Challenge Evaluation 2006) reporting. The ICE iris database contained many very difficult and corrupted images, often in poor focus, and with much eyelid occlusion, with motion blur, raster shear, and sometimes with the iris partly outside of the image frame. As ROC curves require False Matches, NIST used a much more liberal decision criterion than is used in any actual deployments of iris recognition. As seen in Figure 4, using liberal thresholds that generate False Match Rates (FMR) in the range of 0.001–0.00001, score normalization adversely impacts on the ROC curve by increasing the False nonMatch Rate (FnMR). The Equal Error Rate (where FnMR = FMR, indicated by the solid squares) is about 0.001 without score normalization, but 0.002 with the normalization. Similarly at other nominal points of interest in this region of the ROC curve, as tabulated within Figure 4, the cost of score normalization is roughly a doubling in the FnMR, because marginal valid matches are rejected due to the penalty on fewer bits having been available for comparison. In conclusion, whereas Table 1, and Figures 2 and 3 document the important benefit of score normalization when operating with very large databases that require several orders of magnitude higher confidence against False Matches, Figure 4 shows that in scenarios which are much less demanding for FMR, the FnMR is noticeably penalized by score normalization, and so the ROC curve suffers.

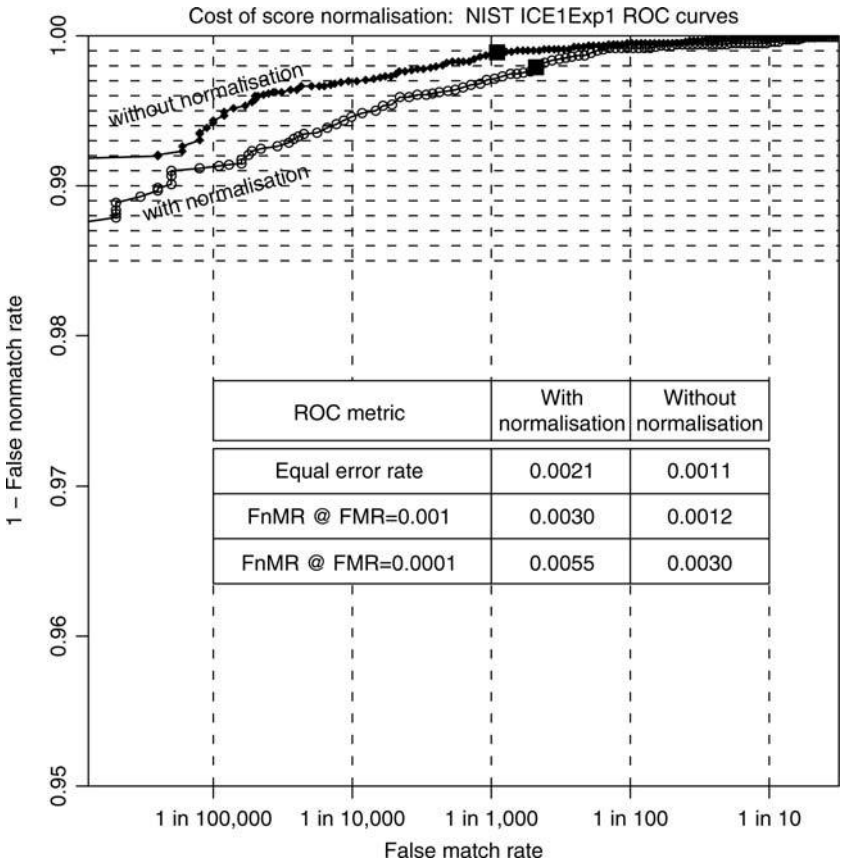
Adapting Decision Thresholds to the Size of a Search Database

Using the SQRT normalization rule, Figure 5 presents a histogram of all 200 billion cross-comparison similarity scores HD_{norm} among the 632,500 different irises in the Middle Eastern database [3]. The vast majority of these IrisCodes from different eyes disagreed in roughly 50% of their bits as expected, since the bits are equiprobable and uncorrelated between different eyes [2, 1]. Very few pairings of IrisCodes could disagree in fewer than 35% or more than 65% of their bits, as is evident from the distribution. The form of this distribution needs to be understood, assuming that it is typical and predictive of any other database, in order to understand how to devise decision rules that compensate for the scale of a search. Without this form of score normalization by the scale of the search, or an adaptive decision threshold rule, False Matches would occur simply because large databases provide so many more opportunities for them.

The solid curve that fits the distribution data very closely in Figure 5 is a binomial probability density function. This theoretical form was chosen because comparisons between bits from different IrisCodes are Bernoulli trials, or conceptually "coin tosses," and Bernoulli trials generate binomial distributions. If one tossed a coin whose probability of "heads" is p in a series of n independent tosses and counted the number m of "heads" outcomes, and if one tallied this fraction $x = m/n$ in many such repeated runs of n tosses, then the expected distribution of x would be as per the solid curve in Figure 5:

$$f(x) = \frac{n!}{m!(n-m)!}\, p^m (1-p)^{(n-m)} \qquad (3)$$

The analogy between tossing coins and comparing bits between different IrisCodes is deep but imperfect, because any given IrisCode has internal correlations arising from iris features, especially in the radial direction [2]. Further correlations are introduced because the patterns are encoded using 2D Gabor wavelet filters, whose lowpass aspect introduces correlations in amplitude, and whose bandpass aspect introduces correlations in phase, both of which linger to an extent that is inversely proportional to the filter bandwidth. The effect of these correlations is to reduce the value of the distribution parameter n to a number significantly smaller than the number of bits that are actually



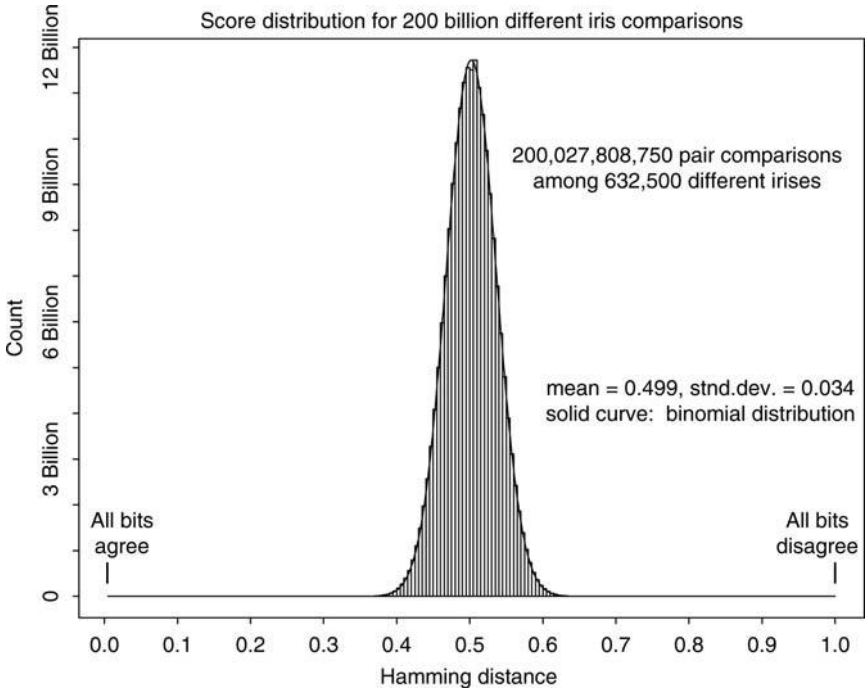
Score Normalization Rules in Iris Recognition. Figure 4 Adverse impact of score normalisation in ROC regions where high False Match Rates are tolerated (e.g., 0.00001 to 0.001 FMR). In these regions, the False nonMatch Rate is roughly doubled as a result of score normalization.

compared between two IrisCodes; n becomes the number of effectively independent bit comparisons. The value of p is very close to 0.5 (empirically 0.499 for this database), because the states of each bit are equiprobable *a priori*, and so any pair of bits from different IrisCodes is equally likely to agree or disagree.

The binomial functional form that describes so well the distribution of normalized similarity scores for comparisons between different iris patterns is key to the robustness of these algorithms in large-scale search applications. The tails of the binomial attenuate extremely rapid, because of the dominating central tendency caused by the factorial terms in (3). Rapidly attenuating tails are critical for a biometric system to survive the vast numbers of opportunities to make False Matches without actually making any, when applied in an "all-against-all" mode of searching for any matching or multiple identities, as is contemplated in some national ID projects. The requirements of biometric operation in "identification" mode by exhaustively searching a large database are vastly more demanding than operating merely in one-to-one "verification" mode (in which an identity must first be explicitly asserted, which is then verified in a yes/no decision by comparison against just the single nominated template).

If P_1 is the False Match probability for single one-to-one verification trials, then $(1-P_1)$ is the probability of not making a False Match in single comparisons. The likelihood of successfully avoiding this in each of N independent attempts is therefore $(1-P_1)^N$, and so P_N, the probability of making at least one False Match when searching a database containing N different patterns is:

$$P_N = 1 - (1 - P_1)^N \tag{4}$$



Score Normalization Rules in Iris Recognition. Figure 5 Binomial distribution of normalised similarity scores in 200 billion comparisons between different eyes. Solid curve is (3).

Observing the approximation that $P_N \approx NP_1$ for small $P_1 << \frac{1}{N} << 1$, when searching a database of size N an identifier needs to be roughly N times better than a verifier to achieve comparable odds against making False Matches. In effect, as the database grows larger and larger, the probability of making a False Match also grows almost in proportion. Obviously the frequency of False Matches over time also increases with the frequency of independent searches that are conducted against the database. In the Middle Eastern deployment [3] from which the data of Figure 5 was taken, in which typically about 12,000 daily arriving passengers are each compared with about a million stored IrisCodes, the total number of iris comparisons is about 12 billion per day. To survive successfully so many opportunities to make False Matches, the decision threshold policy must be adaptive to both of these factors: the size of the database and the frequency of searches through it (the *query rate*). Fortunately, because of the underlying binomial combinatorics, the algorithms with score normalization generate extremely rapidly attenuating tails for the HD_{norm} distribution. The consequence is that extremely small adjustments to tighten the decision threshold yield order-of-magnitude increases in robustness against False Matches, and therefore in large-scale search capability. But before specifying those rules, it is first necesssary to understand the effect of the rotation range over which repeated comparisons are done.

Factoring the Rotation Range into Adaptive Decision Rules

Because cameras or heads may be tilted during iris image acquisition, and indeed the eye itself can undergo torsional (rotational) movements, it is necessary to compare iris patterns at each of several relative orientations before deciding whether they match. Thus, when searching a database of N enrollees and performing each IrisCode comparison in each of k orientations, the total number of comparisons effectively becomes $(k \times N)$. Since the best match (smallest score) found in each set of k comparisons is the score that is retained, the new distribution for comparisons between different eyes is biased towards lower scores, has a lower mean, and has asymmetric tails compared with the unrotated case seen in Figure 5. Using the same database of IrisCodes but now performing all comparisons

in each of $k = 7$ orientations generates the new score distribution seen in Figure 6.

The new distribution after k rotations of all IrisCodes in the search process still has a simple analytic form that can be derived theoretically. Let $f_0(x)$ be the raw density distribution obtained for the HD_{norm} scores between different irises after comparing them only in a single relative orientation; for example, $f_0(x)$ might be the binomial defined in (3). Then $F_0(x)$, the cumulative of $f_0(x)$ from 0 to x, becomes the probability of getting a False Match in such a test when using HD_{norm} acceptance criterion x:

$$F_0(x) = \int_0^x f_0(x)dx \tag{5}$$

or, equivalently,

$$f_0(x) = \frac{d}{dx}F_0(x) \tag{6}$$

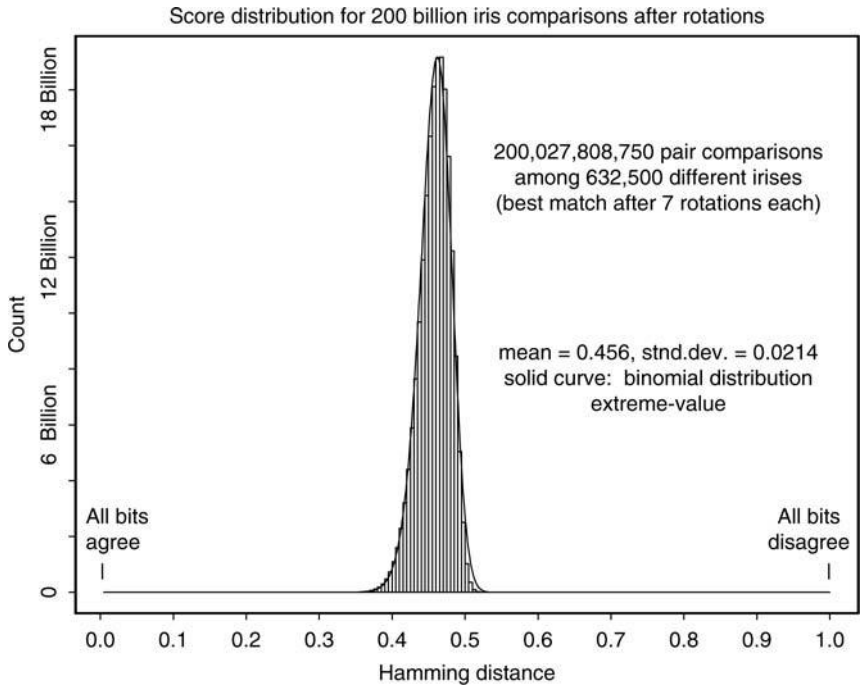
Clearly, then, the probability of *not* making a False Match when using decision criterion x is $1 - F_0(x)$ after a single test, and it is $[1 - F_0(x)]^k$ after carrying out k such tests independently at k different relative orientations. It follows that the probability of a False Match after a "best of k" test of agreement, when using HD_{norm} criterion x, regardless of the actual form of the raw unrotated distribution $f_0(x)$, is:

$$F_k(x) = 1 - [1 - F_0(x)]^k \tag{7}$$

and the expected density $f_k(x)$ associated with this cumulative is:

$$\begin{aligned} f_k(x) &= \frac{d}{dx}F_k(x) \\ &= kf_0(x)[1 - F_0(x)]^{k-1} \end{aligned} \tag{8}$$

It is prudent to allow for at least a 20 *deg* orientation uncertainty, and so $k = 7$ relative rotations in 2.81*deg* intervals were performed when arriving at each best comparison HD_{norm} score plotted in the Figure 6 distribution. As expression (8) in the derivation shown earlier for postrotation probability density



Score Normalization Rules in Iris Recognition. Figure 6 Distribution of normalised similarity scores from the same set of 200 billion iris comparison plotted in Figure 5, but now keeping only the best match after $k = 7$ relative rotations to compensate for unknown degrees of tilt in image acquisition. Solid curve is the $f_k(x)$ density function in (8), using the binomial (3) as the $f_0(x)$ density prior to rotations.

indicates, this means that the left tail of the distribution in Figure 6 is essentially $k = 7$ times higher than the one in Figure 5. The new functional form (8) derived gives an excellent fit to this new distribution, as shown by the solid curve that closely matches the Figure 6 data.

The cumulatives (up to various thresholds) under the left tail of the distribution of normalized similarity scores for different irises compared at $k = 7$ relative tilts, reveal the False Match Rates among the 200 billion iris comparisons if the identification decision policy used each such threshold. These FMR rates are provided in Table 3. Although the smallest observed match was around 0.26, the Table has been extended down to 0.22 using the theoretical cumulative (7) of the extreme value distribution (8) of multiple samples from the binomial (3) plotted as the solid curve in Figure 6, in order to extrapolate the theoretically expected False Match Rates for such decision policies. These False Match Rates, whether empirical or theoretical, also serve as confidence levels that can be associated with a given quality of sample using the score normalization rule (2). In this analysis, only a single eye is presumed to be presented. Under the assumption of independence between right and left eye IrisCodes, which is strongly supported by the available data (see Figure 6 of [2]), the confidence levels in Table 3 could be multiplied together for matches obtained with both eyes.

Now it is finally possible to state a general rule for adaptively selecting a decision criterion threshold on HD_{norm} normalized similarity scores, given the empirical cumulatives shown in Table 3 under the distribution of postrotation scores plotted in Figure 6. If one is performing iris identifications by exhaustive search through an enrolled database of size N, using $k = 7$ relative rotations for every comparison and normalizing raw Hamming Distance scores by the amount of available data as per (2), then the recommended strategy is this:

Calculate the total number of iris comparisons that will be performed in a given period of time, i.e., the size N of the enrolled database times the number of queries against it during that time. Decide the risk tolerance for False Matches during such a period of time, and find the corresponding entry in the second column of Table 3. Then the first column gives the recommended decision threshold on HD_{norm} scores.

Score Normalization Rules in Iris Recognition. Table 3 False match rates with HD_{norm} score normalisation: dependence on criterion (200 Billion cross comparisons)

HD Criterion	Observed False Match Rate
0.220	0 (*theor: 1 in 5* $\times 10^{15}$)
0.225	0 (*theor: 1 in 1* $\times 10^{15}$)
0.230	0 (*theor: 1 in 3* $\times 10^{14}$)
0.235	0 (*theor: 1 in 9* $\times 10^{13}$)
0.240	0 (*theor: 1 in 3* $\times 10^{13}$)
0.245	0 (*theor: 1 in 8* $\times 10^{12}$)
0.250	0 (*theor: 1 in 2* $\times 10^{12}$)
0.255	0 (*theor: 1 in 7* $\times 10^{11}$)
0.262	1 in 200 billion
0.267	1 in 50 billion
0.272	1 in 13 billion
0.277	1 in 2.7 billion
0.282	1 in 284 million
0.287	1 in 96 million
0.292	1 in 40 million
0.297	1 in 18 million
0.302	1 in 8 million
0.307	1 in 4 million
0.312	1 in 2 million
0.317	1 in 1 million

Example: If every month 100,000 passengers are each compared to 1 million enrolled IrisCodes by exhaustive search, generating 10^{11} comparisons per month, and no more than one False Match can be tolerated per month, then Table 3 indicates that the recommended single-eye decision criterion to use would be around 0.265 for HD_{norm} similarity scores.

Related Entries

► Iris Encoding and Recognition using Gabor Wavelets

References

1. Daugman, J.G.: The importance of being random: Statistical principles of iris recognition. Pattern Recognit. **36**, 279–291 (2003)
2. Daugman, J.G.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**, 21–30 (2004)

3. Daugman, J.G.: Probing the uniqueness and randomness of IrisCodes: Results from 200 billion iris pair comparisons. Proc. IEEE **94**(11), 1927–1935 (2006)
4. Daugman, J.G.: New methods in iris recognition. IEEE Trans. Syst. Man Cybern. **37**, 1167–1175 (2007)

Sealed Local Biometric Identity Verification Systems

A sealed local biometric identity verification system is a strong example of a trusted biometric identity verification client. Sealed biometric systems are built around the concept that the entire biometric acquisition and matching process is "sealed" within a demonstrably secure environment (see the following figure). No biometric data can ever enter the system except via the integrated biometric sensor, and no biometric data ever leaves the system, period. In fact there is no way to externally access the biometric data at all. The sealed system securely stores cryptographic credentials for each entity that registers its need to verify the user's identity, and releases those credentials only when that entity has cryptographically identified itself, and the correct user has been biometrically verified to be present.

Figure 1 illustrates the concept of a sealed local biometric identity verification system.

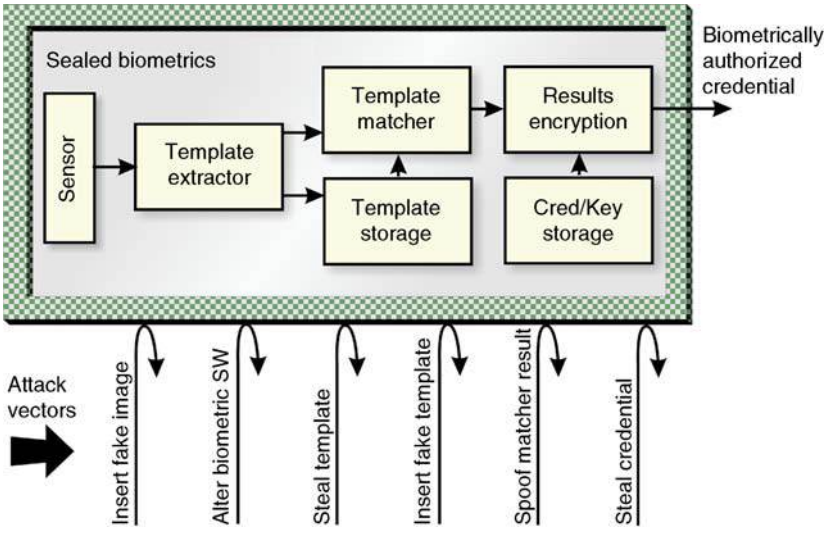
► Biometric Sample Acquisition

Second Level Detail

This refers to the occurrence of fundamental events within the general ridge flow which disturb the regular and parallel flow of ridges. These events can manifest themselves as ending/starting ridges, diverging/merging ridges (bifurcations), and dots and combinations of these events such as short ridges, eyes, spur, and islands. These events are also referred to as Galton characteristics, points or major ridge path deviations.

Minutiae reflect events in a persistent system of papillary ridges that once developed to their final form during gestation remain unchanged throughout life.

Minutiae in the latent keep their basic properties such as relative location, direction, and relations to other points even under adverse conditions (Fig. 1). This differentiates fingerprints from other types of forensic evidence. Minutiae and their formation in



Sealed Local Biometric Identity Verification Systems. Figure 1

sequence are the backbone of fingerprint identification. Additional features that are also referred to as level two are incipient ridges, scars, and creases.

► Fingerprint Matching, Manual

Secure Biometric Token Operating System

► Tamper-proof Operating System

Secure Biometrics

► Biometric and User Data, Binding of

Secure Element

Synonyms

Security processor; Secure token

Definition

A closed, tamper-resistant, well-trusted, and usually certified embedded system with a very lightweight microcontroller and some FLASH storage. Typical secure elements are single-chip entities designed to achieve a black box nature for maximum security. Smart cards, SIM cards, and NFC controllers are the most prevalent examples. Most secure elements interact with a host using a wired or wireless communication channel, but may be standalone systems as well.

► Transportable Asset Protection

Secure Sketch

► Encryption, Biometric

Security and Liveness, Overview

Andy Adler[1], Stephanie Schuckers[2]
[1]Carleton University, Ottawa, ON, Canada
[2]Clarkson University, Potsdam, NY, USA

Definition

The security of a biometric system may be understood to be its resistance to active attacks. Such attacks may be classified as *Presentation attacks (spoofing)*, in which the appearance of the biometric sample is physically changed or replaced; *Biometric processing attacks*, in which an understanding of the biometric algorithm is used to cause incorrect processing and decisions; *Software and networking vulnerabilities* based on attacks against the computer and networks on which the biometric systems run; and *Social and presentation attacks*, in which the authorities using the systems are fooled. This article presents an overview of the techniques used for classifying and assessing these threats. Additionally, newer biometric schemes, such as cancelable biometrics and biometric encryption, that are designed to counter these security threats are reviewed.

Introduction

Security must be defined in the context of an attacker. However, biometric systems, even when not under active attack, should always be assumed to operate in an (at least somewhat) hostile environment – after all, why should one test identity if all can be trusted? The ability of a biometric system to stand up to "zero-effort" attackers is measured by the false accept rate (FAR). Such measures are not typically considered to measure biometric security, but rather biometric performance. This article gives a broad overview of the

security and liveness issues in biometric systems where biometric security is understood to be the resistance of a system to attackers prepared to take active measures to circumvent the system. This article considers four broad types of active attack:

- *Presentation attacks (spoofing):* The appearance of the biometric sample is changed either physically or by replacement with a fabricated sample. For physical biometrics, attackers may then change makeup, facial hair and glasses, or abrade and cut fingerprints in order to avoid being recognized; to attempt to be recognized as another person, a spoofed fingerprint or false iris contact lens may be constructed and placed over the corresponding body part.
- *Biometric processing attacks:* A detailed understanding of the biometric algorithm is used to cause incorrect processing and decisions. The possible attacks depend on the details of the biometric algorithm (► *Biometric vulnerabilities, Overview*). Some examples are enrolling specially crafted noisy images that artificially lower thresholds; regeneration of sample images from stored templates; and ► side-channel attacks based on "leaked" system information such as from match scores or timing of processing.
- *Software and networking vulnerabilities:* Since biometric systems are implemented on server computers, they are vulnerable to all cryptographic, virus, and other attacks, which plague modern computer systems [1]. Examples are databases security, denial of service (DoS) attacks, and overriding the biometric decision with compromised software. These issues are not covered in detail in this article, since they are not unique to biometric system security.
- *Social and presentation attacks:* Security systems depend on a chain of trust. Links in this chain between systems are especially vulnerable [2]. Presentation attacks involve the use of fraudulent identity documents, which may be legitimately issued; social attacks focus on convincing an operator to override or allow fraudulent exceptions. This article points out the importance of these issues like software vulnerabilities, but does not cover them in detail.

This article gives an overview of the security issues in biometric systems, including classifications, security performance measures, liveness and antispoofing and novel biometric protection schemes.

Biometric Security Classifications

Several authors have developed classification schemes, which provide a taxonomy of biometric security challenges. Maltoni et al. [3], classify biometric system vulnerabilities as follows:

- *Circumvention* is an attack by which one gains access to the protected resources by a technical measure to subvert the biometric system. Such an attack may subvert the underlying computer systems (overriding matcher decisions, or replacing database templates) or may involve a replay of valid data.
- *Covert acquisition (contamination)* is the use of biometric information captured from legitimate users to access a system. Examples are spoofing via capture and playback of voice passwords and lifting latent fingerprints to construct a mold. This category can also be considered to cover regenerated biometric images (► *Template Security*). For example, a fingerprint image can be regenerated from the template stored in a database (and these data can be captured covertly [4]). Covert acquisition is worrisome for cross-application usage (eg., biometric records from a ticket for an amusement park used to access bank accounts).
- *Collusion and Coercion* are biometric system vulnerabilities from legitimate system users. The distinction is that, in collusion, the legitimate user is willing (perhaps by bribe), while the coerced user is not (through a physical threat or blackmail). Such vulnerabilities bypass the computer security system, since the biometric features are legitimate. It may be possible to mitigate such threats by automatically detecting the unusual pattern of activity. Such attacks can be mounted from both administrator and user accounts on such a system; attacks from user accounts would first need to perform a privilege escalation attack [1].
- *Denial of Service (DoS)* is an attack that prevents legitimate use of the biometric system. This can take the form of slowing or stopping the system (via an overload of network requests) or by degrading performance. An example of the latter would be enrolling many noisy samples that can make a system automatically decrease its decision threshold and thus increase the FAR. The goal of DoS is often to force a fall back to another system (such as

operator override) that can be more easily circumvented, but DoS may be used for extortion or political reasons.

- *Repudiation* is the case where the attacker denies accessing the system. A corrupt user may deny her actions by claiming that her biometric data were "stolen" (by covert acquisition or circumvention) or that an illegitimate user was able to perform the actions due to the biometric false accept. Interestingly, biometric systems are often presented as a solution to the repudiation problem in the computer security literature [1]. One approach to help prevent repudiation would be to store presented images for later forensic analysis. However, this need must be balanced against user privacy concerns [5].

Another class of biometric vulnerabilities are those faced by the system user, developed by Ratha et al. [6]. These issues impact on the user's privacy and can lead to identity theft or system compromise.

- *Biometrics is not secret:* Technology is readily available to capture images of faces, fingerprints, irises and make recordings of voice or signature – without subject consent or knowledge [2, 7]. From this perspective, biometrics is not secret. On the other hand, from a cryptography or privacy [5] perspective, biometric data are often considered to be private and secret. This distinction is important, as our understanding of computer and network security is centered around the use of secret codes and tokens [1]. For this reason, cryptographic protocols that are not robust against disclosure of biometric samples are flawed.
- *Biometrics cannot be revoked:* A biometric feature is permanently associated with an individual, and a compromised biometric sample will compromise all applications that use that biometric feature. Such compromise may prevent a user from re-enrolling [2]. Note, however, that this concern implies that biometrics is secret, as opposed to the previous consideration. One proposed solution is *Cancelable biometrics*, although the vulnerability of such systems is not well understood.
- *Biometric features have secondary uses:* If an individual uses the same biometric feature in multiple applications, then the user can be tracked if the organizations share data. Another aspect of this problem is *secondary use* of ID cards. For example, a driver's license is designed with the requirements to prove identity and driver certification to a police officer, but it is used to prove age, name and even citizenship. Similarly, biometric applications will be designed with a narrow range of security concerns, but may be used in very different threat environments.

Biometric systems form part of larger security systems and their risks and vulnerabilities must be understood in the context of the larger system requirements. An excellent review of the security of biometric authentication systems is [7]. Each assurance level from "passwords and PINs" to "Hard crypto token" is analyzed to determine which biometric devices are suitable. Since biometric systems are complex and represent many interconnected subsystems, there are many potential points for attack. Vulnerabilities in Biometric Systems are considered in the article ▶ *Biometric Vulnerabilities: Overview*.

Liveness and Spoofing

Clearly, biometric systems are vulnerable to artificial changes to the biometric features. Such changes can be of two types: to *avoid detection* as an enrolled user or watch list candidate and to *masquerade* as another legitimate user. The former is easier and can sometimes be as simple as using glasses, makeup, or abrasions and cuts to the finger. Masquerading or *spoofing* attempts to gain unauthorized access at the biometric sensor with artificial biometric features of authorized users, called "spoofs." This is widely publicized for fingerprint where it is possible to spoof a variety of fingerprint sensors through relatively simple techniques using casts of a finger with molds made of materials, including silicon, Play-Doh, clay, and gelatin (gummy finger). Such spoof molds can be scanned and verified when compared with a live enrolled finger [8–11]. Masquerade is also possible in the scenario of dismembered fingers; cadaver fingers can be scanned and verified against enrolled fingers [9]. It is also possible to spoof other common biometric modalities: for iris and face, using pictures or high resolution video, for iris with contact lenses, with voice with recordings [8, 11].

There are several approaches to increase the difficulty of spoofing: multiple biometric features, liveness, and the use of biometrics in combination with a

challenge response, passwords, tokens, smart cards. The goal of liveness testing is to determine if the biometric feature being captured is an actual measurement from the authorized, live person, who is present at the time of capture. Typically, liveness is a secondary measure after biometric authentication, which must be met in order to achieve a positive response. Liveness and antispoofing methods are covered in detail in the following summaries [12–14]. The need for protections from spoofing may be assessed on an application basis, although there is a need to address the spoofing vulnerability throughout the industry, as the reputation of biometric systems as a security measure must be considered. Liveness detection adds an additional layer of security that also can increase the users trust in biometric technology.

Characteristics of Liveness Approaches: The following characteristics for evaluating biometric systems need to be considered in implementing a liveness algorithm.

- *Ease of use*: liveness approaches vary in the ease of use. For example, a fingerprint deformation approach, which requires a specific rotation procedure, may be considered more difficult to use [15]. A fingerprint sensor using spectroscopy where liveness inherent to the biometric feature may be considered easier to use.
- *Collectability*: liveness approaches vary in the ease of collection. For example, electrocardiogram, which requires two points of contact on opposite sides of the body or pulse oximetry where the finger must be enclosed to protect from ambient light [16].
- *User acceptance*: liveness approaches that may have low user acceptance are teh ones that are more likely to be linked with medical conditions (eg., electrocardiogram, DNA).
- *Universality*: clearly, all authorized users must be live when enrolling. However, the liveness method may be difficult to measure in some subjects. For example, perspiration in fingerprint images may be difficult to measure in subjects with very dry skin.
- *Uniqueness*: For liveness approaches that are inherent to the biometric feature, this is essential. However, it is not clear that, for example, electrocardiogram or gait is unique to large data sets of individuals. Thus, these biometric/liveness approaches may be appropriate for applications with a smaller number of individuals.
- *Permanence* is important to liveness approaches taht are inherent to the biometric feature and where the biometric/liveness features may vary over time. This will impact on the biometric and liveness error rates.
- *Spoof-ability* describes whether the liveness mechanism designed to protect against spoofing can be spoofed. For example, it may be possible to fool pulse oximetry-based liveness, using a clear spoof that allows transmission of the light needed to make the pulse oximetry measurement.

The terms *liveness* and *antispoofing* are not completely synonymous. Measurements that rule out specific spoofs do not absolutely measure liveness. For example, a liveness measure to detect pupil movement will detect attempts based on a simple photograph of a face. However, a modified spoofing method, such as cutting a hole in the picture and putting a real pupil behind it, may result in a successful spoof attempt. Such a spoof is partially alive (to fool the liveness) and partially a spoof (fabricated user biometric feature).

Encoded Biometric Schemes

Classical biometric systems require access to enrolled templates in unencoded form. This differs from traditional computer security systems where a raw password need never be stored. Instead, a cryptographic hash (one-way function) of the password is stored, and each new test password is hashed and compared with the stored version. Since such cryptographic techniques provide important protections, there is great incentive to develop analogous methods for biometric systems. Encoded biometric schemes are designed to avoid these problems by embedding the secret code into the template, in a way that can be decrypted only with an image of the enrolled individual [17, 18]. Since the code is bound to the biometric template, an attacker should not be able to determine either the enrolled biometric image or secret code, even if he had access to the biometric software and hardware. Such technology would enable enhanced privacy

protection, primarily against secondary use of biometric images [5] It would also reduce the vulnerability of network protocols based on biometrics [7]. Biometrically enabled computers and mobile phones currently must hide passwords and keys in software; biometric encryption would protect against this vulnerability. Another interesting application is for control of access to digital content with the aim of preventing copyright infringement. Biometric encryption systems are not widely deployed; research systems still suffer from high error rates and slow processing speed. However, such systems offer some compelling benefits for many applications, and research is active.

Cancelable biometric features (see *Cancelable Biometrics*) are encoded with a distortion scheme that varies from application to application. The concept was developed to address the privacy and security concerns that biometric features are not secret and cannot be canceled. During enrollment, the input biometric image is subjected to a known distortion controlled by a set of parameters. The distorted biometric sample can, in some schemes, be processed with standard biometrics algorithms, which are unaware that the features presented to them are distorted. During matching, the live biometric sample must be distorted with the same parameters, which must be security stored. The cancelable nature of this scheme is provided by the distortion, in that it is not the user's "actual" biometric which is stored, but simply one of an arbitrarily large number of possible permutations. The concern with cancelable biometric features is the security of the storage and transmission of the distortion parameters.

Biometric cryptosystems (► Encryption, Biometric) are designed to overcome many security issues in traditional biometric schemes by avoiding template storage and the match stage of biometric processing. Instead, the biometric features are bound to a secret key that is designed to be recoverable only with a biometric image from the enrolled individual. Clearly, the key difficulty in the design biometric encryption systems is the variability in the biometric image between measurements; the presented biometric image cannot itself be treated as a code, since it varies with each presentation.

The earliest biometric encryption system was proposed by Soutar et al. [18]. Enrollment creates a template binding a secret code to the multiple sample images. During decryption, an error correcting scheme based on Hamming distance is used to allow for variability in the input image. Similar schemes were proposed for voice passwords (in which a vector of features is calculated, and each value is used to select a fraction of the key bits from a table)and iris images.

A significant body of work on biometric encryption has been done in the cryptography community, much based on the *fuzzy vault* construction of Juels and Sudan [19]. This scheme allows a cryptographic encoding with a variable number of un-ordered data points, which makes it suitable for fingerprint minutiae. Clancy et al. [20] designed a fingerprint algorithm that encodes the secret as the coefficients of a Galois field polynomial. Minutiae points are encoded as coordinate pairs, and numerous "chaff" points are added. During key release, the points closest to the new minutiae are chosen, and the key estimated using an error correcting scheme.

Encoded biometric schemes potentially offer some important advantages in security and privacy, since the template does not need to be available in unencrypted form. However, little work has been done to study the security of biometric encryption schemes. Uludag et al. [21] note that most proposed biometric encryption systems only appear to account for a "limited amount of variability in the biometric representation." They suggest that many biometric encryption systems can be attacked simply via the FAR, by presenting biometric samples from a representative population. A cryptographic attack of biometric encryption was developed by Adler [22], based on using any "leaked" information to attempt a hill-climbing of the biometric template. Overall, while biometric encryption offers significant promise, there is little understanding of the practical applicability and security of these systems.

Performance of Biometrics Security and Liveness

In order to quantify and compare the security and liveness performance of biometric systems, it is necessary to have appropriate figures of merit. There exists well understood measures of biometric performance under zero-effort impostor attempts: the false accept (FAR) false reject rates (FRR), failure to acquire, and transaction time among others. It is conceptually reasonable to extend these measures to the active attackers

considered here (although there are clear experimental difficulties in performing the measurements).

In general, a security protection measure is created to protect against a particular active attacker. Using the example of a *liveness (L)* detection system, the following measures are defined:

- *L false reject ratio* (LFRR): the number of times a legitimate attempt is rejected as an attack, divided by the total number of legitimate attempts.
- *L false accept ratio* (LFRR): the number of times an active attack against *L* (a spoof, in the case of liveness) is accepted as legitimate divided by the total number of attack attempts.
- *L failure to acquire*: The number of times the *L* module to is unable to collect information to make a decision, divided by the total number of attempts.
- *L mean transaction time*: The average time required by the *L* module to make a decision.

In a general biometric system, one or more security protection measures (*L*) will function in addition to the core biometric (*B*) decisions. The performance of the combination of a security measure and a biometric matcher is defined as the *combined system performance*, with the following measures:

- *System false reject ratio* (SFRR): the number of times a legitimate attempt is rejected as an attack (by *L*) or an impostor (by *B*), divided by the total number of legitimate attempts. Here, false rejects are the combined set of errors from the biometric stage (false reject of the correct person) and errors from the liveness stage (*L* false reject). Thus, the SFRR is the union of the FRR and the LFRR. In general, SFRR $\leq$ FRR + LFRR, because some transactions may be rejected by both *L* and *B*.
- *System false accept ratio* (SFAR): the number of times an active attack or an impostor is accepted as legitimate divided by the total number of attack or impostor attempts. This definition is more complicated, since the measure must combine evaluations of spoof accepts (against *L*) and traditional false accepts (against *B*). This measure is modified by the expected frequency of impostor and attack attempts, and thus by the relative weight of these events in the test database.
- *System failure to acquire ratio*: The number of times the *L* module or the biometric system *B* is unable collect information to make a decision divided by the total number of attempts.
- *System mean transaction time*: The average time required by the entire system to make all decisions (including the liveness and match decisions).

Clearly, the main difficulty in making these measurements is developing a database or procedures for the active attacks, which are somehow reflective of their expected frequency in the target operational conditions. Nevertheless, such measures are important to clarify how security measures impact on the overall system performance. For example, a biometric system with very good performance (1% EER) will be greatly impacted by a liveness algorithm that has a liveness equal error rate of 5%. In this case, the system false reject ratio is equal to the union of the two measures, that is between 5% and 6%. This would represent a dramatically worse system in terms of the experience of its users.

Summary

The security of a biometric system is its resistance to active attack. Such attacks may be classified as *Presentation attacks (spoofing)*, in which the appearance of the biometric sample is physically changed or replaced; *Biometric processing attacks*, in which an understanding of the biometric algorithm is used to cause incorrect processing and decisions; *Software and networking vulnerabilities*, based on attacks against the computer and networks on which the biometric systems run; and *Social and presentation attacks*, in which the authorities using the systems are fooled. In this article, a survey of issues in biometric security and liveness (anti-spoofing) have been presented, including frameworks to classify and measure biometric security performance. In addition, encoded biometric schemes are reviewed to clarify their promise to to counter these security threats. Overall, in the design of security and liveness systems, it is important to consider the operational requirements of the application and the specific security threats against which it will be tested.

Related Entries

► Encryption, Biometric
► Biometric Security, Overview

- ▶ Biometric System Design
- ▶ Biometrics and Security, Standardization
- ▶ Biometric Vulnerabilities, Overview
- ▶ Cancelable Biometrics
- ▶ Fraud reduction
- ▶ Liveness Detection: Face
- ▶ Liveness Detection: Fingerprint
- ▶ Liveness Detection: Iris
- ▶ Liveness Detection: Voice
- ▶ User Interface, System Design
- ▶ Zero-effort Forgery Test

References

1. Ferguson, N., Schneier, B.: Practical Cryptography. John Wiley & Sons, NJ, USA (2003)
2. Schneier, B.: The Uses and Abuses of Biometrics. Communications of the ACM vol. 42, pp. 136 (1999)
3. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, Berlin (2003)
4. The Guardian (17 Nov. 2006) Cracked it!
5. Cavoukian, A.: Privacy and Biometrics, In Proceedings of the Intertnational Conference on Privacy and Personal Data Protection, Hong Kong, China (1999)
6. Ratha, N.K., Connell, J.H., Bolle, R.M.: Enhancing security and privacy in biometrics-based authentication systems. IBM Syst. J. **40**, 614–634 (2001)
7. International Committee for Information Technology Standards (INCITS) Study Report on Biometrics in E-Authentication, Technical Report INCITS M1/06-0693 (2006)
8. Thalheim, L., Krissler, J.: Body Check: Biometric Access Protection Devices and their Programs Put to the Test, ct magazine, November (2002)
9. Schuckers, S.A.C.: Spoofing and anti-spoofing measures, Information Security Technical Report, vol. 7, pp. 56–62 (2002)
10. International Biometric Group, (2007) Spoof, Test underway of fingerprint and iris recognition systems' resistance to spoofing http://www.biometricgroup.com/spoof/ 2007. Accessed 7 April, 2009
11. Matsumoto, T.: Gummy Finger and Paper Iris: An Update, Workshop Inform Security Research, Fukuoka, Japan, October (2004)
12. International Biometric Group, Liveness Detection in Biometric Systems, http://www.ibgweb.com/reports/public/reports/liveness.html
13. Schuckers, S.A.C., Derakhshani, R., Parthasardhi, S., Hornak, L.A., (2006) Liveness Detection in Biometric Devices, in Electrical Engineering Handbook, 3rd edition, CRC Press, Chapter 26, ISBN: 084932274X
14. Coli, P., Marcialis, G.L., Roli, F.: Vitality Detection from Fingerprint Images: A Critical Survey, Advances in Biometrics, Springerlink, vol. 4642, pp. 722–731 (2007)
15. Antonelli, A., Cappelli, R., Maio, D., Maltoni, D.: Fake Finger Detection by Skin Distortion Analysis. IEEE Trans. Inf. Forensics Secur. **1**(3), 360–373 (2006)
16. Biel, L., Pettersson, O., Philipson, L., Wide, P.: ECG analysis: A new approach in human identification. IEEE Trans. Instrum Meas. **50**, 808–812 (2001)
17. Davida, G.I., Frankel, Y., Matt, B.J.: On enabling secure applications through off-line biometric identification. In Proceedings of the IEEE Symposium on the Privacy and Security. pp. 148–157 (1998)
18. Soutar, C., Roberge, D., Stoianov, A., Gilroy, R., Vijaya Kumar, B.: Biometric Encryption using image processing. In Proc. SPIE Int. Soc. Opt. Eng. **3314**, 178–188 (1998)
19. Juels, A., Sudan, M.: A fuzzy vault scheme, In Proceedings of the IEEE International Symposium on the Information Theory. pp. 408 (2002)
20. Clancy, T.C., Kiyavash, N., Lin, D.J.: Secure smartcard-based fingerprint authentication. In Proceedings of the ACMSIGMM 2003 Multimedia, Biometrics Methods and Applications Workshop. pp. 45–52 (2003)
21. Uludag, U., Pankanti, S., Prabhakar, S., Jain, A.K.: Biometric Cryptosystems: Issues and Challenges. Proceedings of the IEEE. vol. 92, pp. 948–960 (2004)
22. Adler, A.: Vulnerabilities in biometric encryption systems. In Proceedings of the AVBPA, Tarrytown, NY, USA, LNCS 3546: 1100–1109 (2005)

Security Block

- ▶ Common Biometric Exchange Formats Framework Standardization

Security Issues, System Design

Karthik Nandakumar
Institute for Infocomm Research A *STAR, Fusionopolis, Singapore

Definition

Person authentication is one of the critical tasks in a securing information technology (IT) systems and biometric recognition is a natural and reliable solution that can provide secure authentication. However, a biometric system is just one component of the overall IT security solution. To ensure the confidentiality

of the biometric information and the integrity of the biometric system, several security issues must be addressed in the design stage. Appropriate steps must be taken to guard against the vulnerabilities at the interfaces between the different components of the security system and the threats introduced due to improper implementation and administration of the biometric system. Furthermore, the security of a biometric system must be analyzed systematically based on standard methodologies such as the Common Criteria framework.

Introduction

In today's digital world, a wide variety of information technology (IT) systems is used by the government (e.g., e-governance) and private organizations (e.g., e-commerce) to deliver their products and services to the society. The security of these IT systems is of vital importance because any security breach could potentially lead to adverse consequences such as terrorist attacks, financial frauds and loss of privacy. In the context of IT systems, security can be defined as "protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction" [1]. In general, there are four major aspects to be considered in IT system security.

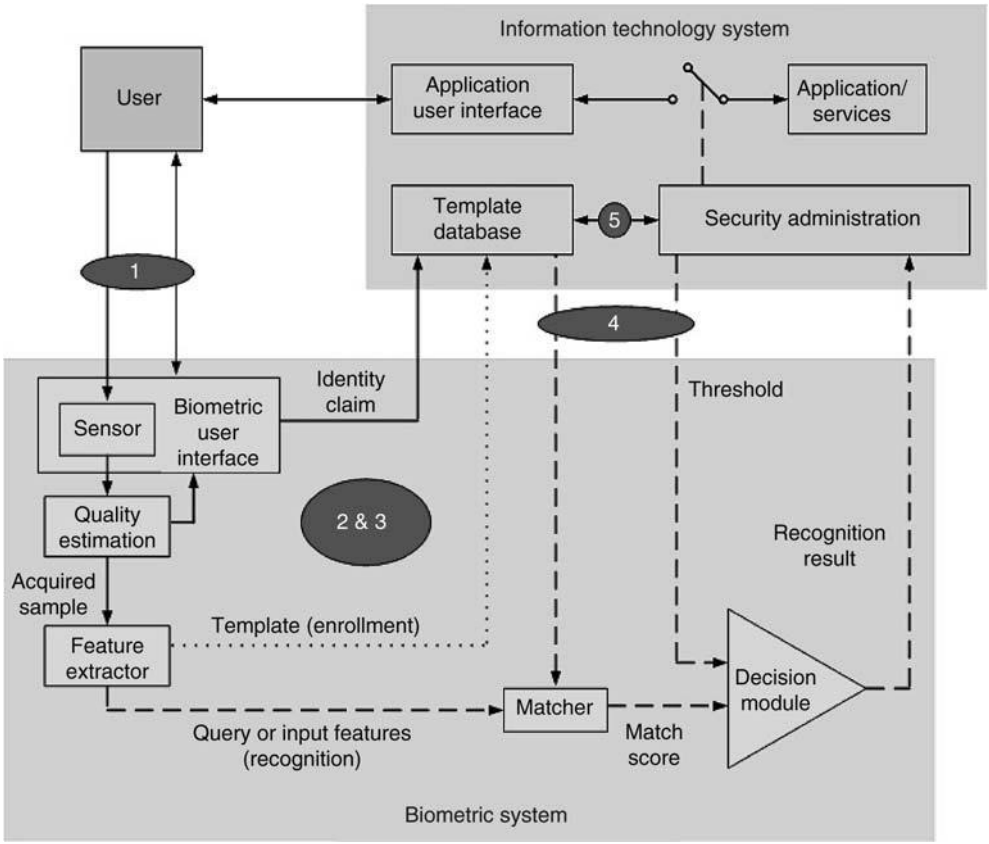
- *Secure authentication* – only legitimate/authorized users should able to access the system and carry out specific tasks.
- *Data confidentiality* – prevent illegitimate access or disclosure of sensitive data or information.
- *Integrity* – guard against improper modification or destruction of the system/data and ensure nonrepudiation and authenticity of information.
- *Availability* – guarantee timely and reliable access to and use of information.

Biometric recognition is one of the techniques that can effectively address the secure authentication problem. Since biometric traits cannot be easily lost, stolen, misplaced or shared, biometric recognition offers a natural and more reliable authentication solution compared to other techniques such as passwords or physical tokens (e.g., ID cards). However, it is important to realize that a biometric system is just one component of the overall IT security solution because it address only the secure authentication aspect. Other technologies such as encryption, digital signature, etc. are needed to meet the confidentiality, integrity, and availability requirements of the complete IT system. Moreover, the biometric system can itself be considered as an independent subsystem within the complete IT system. If the biometric system (or subsystem) is compromised or circumvented, the security of the entire IT system gets affected. Due to this reason, the security aspects involved in the design and implementation of a biometric system should be analyzed independently [2–4].

Biometric System Architecture

The general architecture of a biometric system is shown in Fig. 1. There are four major modules in a generic biometric system, namely, sensor, feature extractor, matcher, and decision module. Sensor is a part of the interface between the user and the biometric system and its function is to scan the biometric trait of the user. The biometric user interface may also have other functionalities such as collecting the identity claim from the user and providing feedback or guidance to the user on how to present the biometric data. Feature extraction module processes the scanned biometric data to extract the salient information (feature set) that is useful in distinguishing between different users. During enrollment, the extracted feature set is stored in a database as a template, which is generally indexed by the user's identity information. The template database is usually considered as a component of the complete IT system. The creation and maintenance of the template database is handled by the security administrator of the IT system. The matcher module is usually an executable program, which accepts two biometric feature sets (one from the template and the other from the query) as inputs, and outputs a match score indicating the similarity between the two sets. Finally, the decision module makes the identity decision based on the threshold set by the system administrator. Based on the recognition results, the user is granted or refused access to the IT application or service by the administrator.

From Fig. 1, it is clear that there are five main vulnerable areas in a biometric system, where security is of critical importance. These areas include (1) the interface between the user and the biometric system,



Security Issues, System Design. **Figure 1** General architecture of a biometric system. The major areas of vulnerability are (1) user-biometric system interface, (2) biometric system modules, (3) interconnections between biometric modules, (4) biometric system-IT system interface and (5) security administration.

(2) the four modules of the biometric system, (3) the interconnection between the modules within the biometric system, (4) the interface between the biometric system and the IT system, and (5) the administration of the template database. The various threats and countermeasures that need to be considered in the design of a biometric system are presented in the subsequent sections. The threat agents (attackers) could either be insiders (authorized users, operators or administrators) or external adversaries. For insider threats, the cause for the threat could be (1) unintentional (inadvertent) error, (2) collusion between attackers, and (3) coercion by external adversaries.

Threats at the Interface between User and Biometric System

In general, any attempt by an attacker to break into the system by presenting a biometric trait can be considered as a threat at the user-biometric system interface level. At this level, the following threats and countermeasures are possible.

- *Casual impersonation*: An impostor attempts to fool the system by presenting his/her biometric trait and impersonating as an authorized user. In this case, the identity to be attacked is chosen randomly and the impostor does not modify his/her own biometric identifiers in any way. The probability of success in such an attack is measured by the false accept rate (FAR) of the biometric system. Since this threat is due to the ▶ intrinsic failure of the biometric system and does not require any explicit effort by the attacker, it is referred to as *zero-effort attack*. This attack can be countered by selecting a very low value of FAR and by restricting the number of failure attempts allowed within a time-frame.
- *Targeted impersonation*: Same as casual impersonation, except that the impostor attacks a specific

identity, which is known to be easier to impersonate (also known as lambs [5]). This attack exploits the fact that FAR is not uniform across all users. The impostor may also target an identity whose biometric characteristics are known to be similar to his/her traits (also known as "Evil Twin" [6]). The same countermeasures used against casual impersonation may be employed to limit the success of this attack.

- *Mimicry*: The impostor may be able to modify his biometric characteristics to match that of the identity under attacks. Examples of this attack include changing one's voice, forging a signature or mimicking a gait pattern. This threat is more common in systems using behavioral biometric traits and in unattended applications. Countering this attack requires biometric systems that have low FAR under skilled forgery.
- *Spoofing*: This is the most common attack at user interface level and it involves the presentation of a spoof (fake or artificial) biometric trait (e.g., dummy finger, recorded voice, etc.). If the sensor is unable to distinguish between fake and genuine biometric traits, the adversary easily intrudes the system under a false identity. This attack requires knowledge of the biometric trait corresponding to the identity to be attacked. This knowledge could be obtained in one of the following three ways: (1) directly colluding with or coercing an authorized user, (2) covert acquisition (e.g., lifting residual fingerprint impressions covertly from the sensor or any surface touched by the authorized user) and (3) stealing the biometric template from a database and reverse engineering the template. The solution to counter this threat is to incorporate liveness detection capability in the biometric sensor. A number of efforts have been made in developing hardware as well as software solutions that are capable of performing liveness detection.
- *Presentation of poor image*: This threat is mainly applicable to screening applications, where the attacker may attempt to hide his true identity by presenting a poor image or noisy biometric sample that may not be matched to his/her template in the database. However, it may also be applicable in verification systems that employ a fall-back mechanism to handle false rejects. In this scenario, the impostor may attempt to bypass the biometric system by providing noisy samples and exploit the loopholes in the exception procedures. This attack can be countered by configuring the biometric system in such a way that the False Nonmatch Rate (FNMR) or False Reject Rate (FRR) is very low.
- *Illegal enrollment*: The impostor may enroll himself into the system illegally (under a false identity) by producing his biometric traits along with false credentials (e.g., fake passports, birth certificates, etc.). It must be emphasized that the secure authentication functionality provided by a biometric system is only as good as the integrity of the enrollment process.

Threats at the Biometric System Modules

Though the sensor, feature extractor, matcher, and decision modules logically constitute a single unit (known as the biometric system or device) within the IT system, there is a variety of possibilities in the physical configuration of these modules. For example, it is possible to place all the four modules and the interfaces between them on a single smart card (or more generally a secure processor). In such systems, known as system-on-card technology, sensor, feature extractor, matcher, and even the templates reside on the card or the chip [7]. The advantage of this technology is that the biometric information never leaves the card and only the recognition results are transmitted to the IT system. It is much easier to design a trusted or secure biometric system based on the system-on-card technology. On the other extreme, consider a large Automated Fingerprint Identification System (AFIS) used in forensic applications. In the AFIS scenario, the modules of the biometric system are typically distributed across different physical locations (sensor may be at the crime scene, feature extractor and decision module may be at the regional investigation office, matcher and database at a national center, etc.). Other intermediate configurations where the sensor and feature extractor may reside together at a remote location, while the matcher and database reside on the server (or a smart card (match-on-card technology)) are also possible. Despite the wide ranging physical configurations, four common attacks can be mounted on the biometric system modules.

- *Bypass*: The attacker may completely bypass one or more modules of the biometric system. For instance, the attacker can bypass the sensor and present the biometric image directly to the feature extractor. In cryptography, this threat is known as

"man-in-the middle" attack. One method to overcome this threat is to employ a trusted biometric system. A trusted biometric system is one in which the different modules are bound together physically and/or logically using ► mutual authentication between the modules.

- *Modification*: The executable program at a module can be modified such that it always outputs the values desired by the adversary. Such attacks are also known as Trojan-horse attacks. Secure code execution practices or specialized tamper-resistant hardware that can enforce secure execution of software can be used to avoid modification of the module functionalities.
- *Exploitation of vulnerabilities*: The attacker may identify and exploit the loopholes in the implementation of the biometric algorithms or insecure configuration to circumvent the biometric system. As an example, consider a matching module in which a specific input value, say X_0, is not handled appropriately and whenever X_0 is input to the matcher, it always outputs a match (accept) decision. This vulnerability might not affect the normal functioning of the system because the probability of X_0 being generated from a real-biometric data may be negligible. However, an adversary can exploit this loophole to easily breach the security without being detected. Note that the attacker may need to bypass one or more modules in the biometric system to exploit such vulnerabilities.
- *Sabotage*: Finally, an adversary can physically tamper with or damage the infrastructure of a biometric system thereby preventing legitimate users from accessing the application. Examples of sabotage include disabling the power supply, damaging the sensor surface or introducing excessive noise (interference) that prevents the normal operation of the system. Apart from causing denial of service to authorized users, this kind of attack may also be used to gain unauthorized access by exploiting the vulnerabilities in the fall-back system.

Threats at the Interconnections between Biometric Modules

The following vulnerabilities are possible when an adversary gains control of the communication interfaces between different modules of the biometric system. Juels et al. [8] outlined the security and privacy issues introduced by insecure communication channels in an e-passport application that uses biometric recognition.

- *Replay attacks*: If the channels between the biometric modules are not secured physically or cryptographically, an adversary may intercept the data being transferred and replay it at a later time. The raw biometric data or extracted features can be intercepted and replayed. Replay attacks are possible even if the data are encrypted. A countermeasure against this attack is to use time-stamps or a challenge/response mechanism. Mutual authentication between the modules and use of one-time session keys during every transaction could also mitigate replay attacks.
- *Hill Climbing attacks*: Hill climbing attacks are possible when the adversary has the ability to inject raw biometric data or features directly into the channel by bypassing the sensor or feature extractor (or through Trojan-horse attacks). This attack also requires some feedback from the biometric system such as the match score [9]. In this scenario, an artificially generated biometric sample or feature set is first introduced into the system and the response (match score) is noted. The adversary then perturbs the initial sample or feature set, submits it to the system and records the new match score. If the match score in the second iteration is higher than the first one, the changes are retained; else, they are discarded. This process is iterated several times until the match score crosses the threshold set by the system administrator. In each iteration where the match score is higher than before, the artificially generated sample or feature set becomes more similar to the template that is being targeted. Restricting the number of failure attempts allowed within a time-frame, increasing the granularity of the match score, the use of trusted biometric systems, etc. are some techniques that can counter the threat of a hill-climbing attack.
- *Sabotage*: The adversary can also sabotage the biometric system by physically damaging the communication interfaces between the modules or by place an interfering source near the communication channel (e.g., a jammer to obstruct a wireless interface). This may cause the denial of service or lead to unauthorized access attempts that try to exploit vulnerabilities in the fall-back system.

Threats at the Interface between Biometric and IT systems

The communication channels at the interface between the biometric system and the IT system carry four main types of information, namely, the user identity information, the biometric template, the system parameters such as threshold, and the recognition results. Since all the four pieces of information are of vital importance, careful attention needs to be given to secure these links. As in the interfaces between the modules, the key threat is the interception and/or modification of data. For instance, if the recognition results can be intercepted and modified by the adversary, the complete biometric system gets bypassed and the security provided by the biometric system is rendered useless. The countermeasures against this type of threat are essentially the same as those used to secure the links between the biometric modules.

Administration Threats

Administration attack refers to all vulnerabilities introduced due to improper administration of the biometric system, which may occur due to the following causes.

- *Insider threat:* Authorized users or system administrators may exceed their authority either inadvertently or with malicious intent. Steps such as security awareness training and audit trails can minimize the threats posed by the insider attacks. However, it is important to ensure the integrity of the audit logs themselves, because any unauthorized tampering of audit logs can lead to undetected insider attacks.
- *Template Modification:* The template database could be hacked or modified by an adversary to gain unauthorized access. This scenario is also applicable to the case where the template is stored on a smart-card. The card may be forged to contain the biometric template of the impostor. Template protection approaches such as encryption and ▶ biometric cryptosystems can be used to prevent this attack.
- *Leakage of Biometric Information:* Leakage of the biometric template information may lead to the following consequences: (1) A physical spoof can be created from the template (see [10, 11]) to gain unauthorized access to the system (as well as other systems that use the same biometric trait), (2) the stolen template can be replayed to the matcher to gain unauthorized access, and (3) the templates can be used for cross-matching across different databases to covertly track a person without his/her consent. Template protection schemes like encryption, feature transformation, and biometric cryptosystems can be used to mitigate this threat.
- *Exception processing*: User authentication systems are usually riddled with exception processing procedures (or fall-back systems) to avoid inconvenience to genuine users. For example, when a user suffers an injury to his finger, he may still be granted access based on alternative authentication mechanisms without undergoing fingerprint recognition. Such exception processing procedures can be easily abused to circumvent a biometric system.

Common Criteria Framework

The Common Criteria (CC) for Information Technology Security Evaluation is an international standard (ISO/IEC 15408) for evaluating the security of information technology systems [12]. This standard provides the ability to compare independent security evaluations of a product or system. The IT product or system whose security properties are to be analyzed is called the Target of Evaluation (ToE). The CC framework classifies the IT security requirements into security functionalities and assurance levels. It also defines two ways of expressing these requirements, namely, Protection Profile (PP) and Security Target (ST). The protection profile (PP) can be used by consumers and system designers to list the various threats and vulnerabilities faced by the ToE and the desired security features that will meet their needs. The security target (ST) precisely specifies the security capabilities of the ToE and the ST is used by the evaluators as the basis for the security evaluation. The ST may also claim compliance with a protection profile. The CC framework can be used to systematically evaluate the security of a biometric system. For example, biometric protection profiles have been introduced in the U.K., U.S. and Germany [6, 13, 14].

Summary

A biometric system is one of the key components in an IT security system that provides the secure authentication functionality. However, the biometric system itself

is vulnerable to a number of security threats and a systematic analysis of these threats is essential when designing a biometric system. In this article, a high-level categorization of the various vulnerabilities of a biometric system was presented and countermeasures that have been proposed to address these threats were discussed. Public acceptance of biometric recognition technology will depend on the ability of system designers to demonstrate that these systems are robust, have low error rates, and are tamper-proof. This can be achieved by evaluating the biometric system security using IT security standards such as the Common Criteria framework.

Related Entries

- ▶ Biometric Matcher
- ▶ Biometric Vulnerabilities, Overview
- ▶ Feature Extraction
- ▶ Security and Liveness: Overview
- ▶ Sensor
- ▶ Template Security

References

1. Coordination of Federal Information Policy – Information Security. In: United States Code, 3532, chap. 35. Available at http://www.law.cornell.edu/uscode/html/uscode44/usc_sec_44_00003532——000-.html
2. Wirtz, B.: Biometric System Security – Part 1. Biometric Technology Today pp. 6–8 (2003)
3. Wirtz, B.: Biometric System Security – Part 2. Biometric Technology Today pp. 8–9 (2003)
4. Soutar, C.: Automatic Fingerprint Recognition Systems. Springer, Berlin (2004)
5. Doddington, G., Liggett, W., Martin, A., Przybocki, M., Reynolds, D.: Sheep, Goats, Lambs and Wolves: A Statistical Analysis of Speaker Performance in the NIST 1998 Speaker Recognition Evaluation. In: Proceedings of the Fifth International Conference on Spoken Language Processing (ICSLP). Sydney, Australia (1998)
6. UK Government Biometrics Working Group: Biometric Device Protection Profile (BDPP). Technical Report Draft Issue 0.82 (2001)
7. Jain, A.K., Pankanti, S.: A Touch of Money. IEEE Spectrum **3**(7), 22–27 (2006)
8. Juels, A., Molnar, D., Wagner, D.: Security and Privacy Issues in E-passports. In: Proceedings of First International Conference on Security and Privacy for Emerging Areas in Communications Networks, pp. 74–88. Athens, Greece (2005)
9. Adler, A.: Sample images can be independently restored from face recognition templates. In: Proceedings of Canadian Conference on Electrical and Computer Engineering, vol. 2, pp. 1163–1166. Montreal, Canada (2003)
10. Ross, A., Shah, J., Jain, A.K.: From Templates to Images: Reconstructing Fingerprints From Minutiae Points. IEEE Trans. Pattern Analy. Mach. Intell. **29**(4), 544–560 (2007)
11. Cappelli, R., Lumini, A., Maio, D., Maltoni, D.: Fingerprint Image Reconstruction From Standard Templates. IEEE Trans. Pattern Analy. Mach. Intell. **29**(9), 1489–1503 (2007)
12. Evaluation Criteria for IT Security. In: ISO/IEC 15408-1. International Standards Organization (2005)
13. Information Assurance Directorate: U.S. Government Biometric Verification Mode Protection Profile for Basic Robustness Environments. Technical Report 1.1 (2007)
14. Security Evaluation of Biometrics. In: ISO/IEC CD 19792. International Standards Organization (2008)

Security Threat Assessment

Security threat assessment is the use of available information, including biographic and biometric data, to perform of a series of investigations to determine whether an individual poses (or is suspected of posing) a threat to transportation or national security. It may include name-based checks, criminal history background checks, and watch list checks.

- ▶ Registered Traveler

Segmentation of Iris Images from Noncooperative Subjects

- ▶ Segmentation of Off-Axis Iris Images

Segmentation of Off-Axis Iris Images

Lauren R. Kennell, Ryan N. Rakvic, Randy P. Broussard
The John Hopkins University, Applied Physics Laboratory, Maryland, USA

Synonyms

Off-angle *or* Nonorthogonal segmentation; Segmentation of iris images from noncooperative subjects.

Definition

Iris image *segmentation* is the process of finding the iris region within an image of the subject's eye area: distinguishing the iris ring from the sclera, eyelids, and other regions, and also detecting obscurations caused by eyelashes and specular reflection. *Off-axis* images are those captured when the subject is not looking in the direction of the camera.

Segmentation of Off-Axis Iris Images. Figure 1 Off-axis iris image from [1], Fig. 9a, p. 586. [Copyright U.S Naval Academy]

Introduction

Iris recognition systems typically require that users position themselves a few inches away from the camera and look into the camera for a few seconds, known as the "stop and stare." More forgiving and user-friendly interfaces in commercial systems may facilitate the introduction of iris systems into a wider range of applications. Taking it a step further, eliminating the user cooperation requirement entirely would enable deployment of systems in which the subject may not be aware of the image capture, such as in a surveillance situation. The less user cooperation that is expected, the more the systems depend on recognition from *off-axis* images, namely those taken when the subject is not looking in the direction of the camera, such as in Fig. 1. The topic of this chapter is the segmentation of off-axis iris images: the task of determining which pixels from the captured image belong to the iris, as opposed to the sclera, the eyelids, eyelashes, etc.

Segmentation methods can incorporate certain parameters, regarding the approximate size of the iris as imaged through a given camera such as the approximate pixel diameter. In addition, segmentation often exploits the basic shape and proportions of a typical eye, but to different degrees and in different ways. Often the pupil (inner) boundary and limbic (outer) boundary are conceptualized as nearly-concentric circles (alternatively, the pupil mass may be thought of as a disk and the limbic boundary as a circle), so one often proceeds by searching for these shapes, for instance with circular edge detection in the appropriate diameter range, and then refining the search as needed to account for irregularly shaped boundaries. Generally speaking, this is still true in the off-axis case, except that the circular model must now be more flexible, replaced with something like an oriented ellipse, to account for sideward and/or up- or downward gaze (yaw and pitch). An off-axis segmentation procedure is hopefully rather robust, because reduced user cooperation may result in lower image quality, such as from blur, inconsistent illumination and shadows, significant eyelid and eyelash obscuration, and specular reflections intersecting the pupil or limbic boundary.

Shape models may enhance robustness and save some time by narrowing down what to look for, and they are useful for other reasons too. A shape model of the inner and/or outer boundary, such as an ellipse with a certain center, axis lengths, and orientation, provides a means to determine the gaze angle, e.g., using eccentricity-based computations, from which the eye image may be projected to orthographic (on-axis) view before feature encoding. Whatever the assumed shape of the iris ring, in many on-axis recognition algorithms the iris boundaries are often identified as closed curves (projecting behind any occluding eyelids), because this description of the boundaries is used to "cut and unwrap" the iris ring into a rectangle prior to feature extraction.

What follows is a summary of off-axis segmentation techniques, many of them adapted from or inspired by previous on-axis methods – clever constructions and combinations of morphology, contours, filters, snakes, splines, trigonometry, transforms, and plain brute force.

Discussion of Segmentation Methods

The Daugman Algorithm

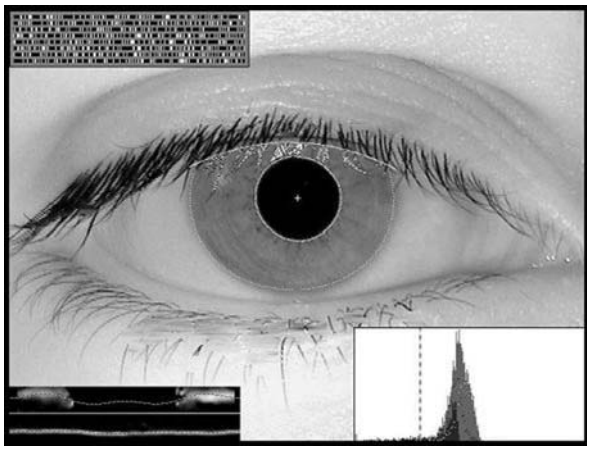
The algorithm developed by Daugman [2–5] is used in the current public deployments of iris recognition systems. Indeed, two of the iris image preprocessing steps introduced in [2], an integro-differential operator for edge detection and the pseudo-polar coordinate transform which transforms (or "unwraps") the iris ring into a doubly-dimensionless coordinate system, have been incorporated into various other proposed recognition methods. Therefore it is natural to begin with the off-axis segmentation in [5].

The segmentation begins with approximating the pupil, limbic, and eyelid boundaries using the integro-differential operator from [2]:

$$\max_{(r,x_0,y_0)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r}\, ds \right|. \qquad (1)$$

Here $I(x, y)$ are the image grayscale values, $G_\sigma(r)$ is a smoothing function such as a Gaussian of scale σ, and the contour integral is along circles given by center (x_0, y_0) and radius r (but the circles are replaced by arcs when searching for eyelids). Thus this operator produces initial circular estimates for the pupil and limbic boundaries from the maximum blurred partial derivative of the image grayscale values with respect to a radial variable. To zero in on the actual boundary, the plots of radial gradients for the pupil and limbic boundaries form two "snakes", like those shown in Fig. 2, each of which is then approximated by a discrete Fourier series. The iris ring is thus bounded by smooth closed curves which would project behind occluding eyelids, but in general the inner and outer boundary will be neither exactly circular nor elliptical.

The pupil boundary contour provides the information necessary to compute the gaze angle from "Fourier-based trigonometry": using the Fourier coefficients to fit an oriented ellipse to the pupil boundary. An ellipse is parameterized by the equations $x(t) = A\cos(t)$, $y(t) = B\sin(t)$, where $A < B$ in the case of a vertical major axis. If the gaze is deviated upward or downward, the pupil boundary would appear rotated by some angle θ, so the more general parameterization is an oriented ellipse, given by

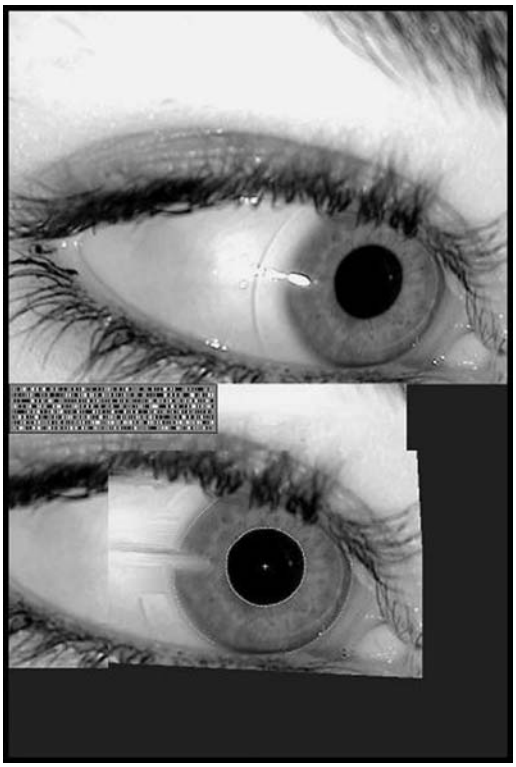


Segmentation of Off-Axis Iris Images. Figure 2
Reprinted from [5], Fig. 4, pp. 1171. Iris image (ICE-1 file 239766) overlaid with pupil, iris-sclera, and eyelid boundaries, and detected eyelashes. Also shows iris pixel histograms (lower right), snakes (the two lower left plots), and resulting Iris Code (top left). In the top snake plot the two fuzzy bands are the curvature map for the iris-sclera boundary, and the dotted curve shows its Fourier series approximation are for the pupil boundary. Copyright 2007 IEEE.

$$\begin{aligned} X(t) &= [A\cos^2\theta + B\sin^2\theta]\cos(t) \\ &\quad + [(B-A)\cos\theta\sin\theta]\sin(t), \\ Y(t) &= [(B-A)\cos\theta\sin\theta]\cos(t) \\ &\quad + [B\cos^2\theta + A\sin^2\theta]\sin(t). \end{aligned}$$

By comparison of these expressions with the lower frequency terms of the Fourier series which describes the pupil boundary contour, the values of A, B, and θ are uniquely determined. This is sufficient to determine sideward and upward/downward gaze directions, and to rotate the image back to orthographic view with an affine transformation. An example image before and after rotation is shown in Fig. 3.

The last step in the segmentation process is the detection of eyelashes, which are found by the realization that the presence of eyelashes overlying the iris results in too many dark pixels in the upper half of the iris, compared with the grayscale distribution in the lower half of the iris. In such cases, when the grayscale distribution in the top half of the iris shows multimodal mixing, the suspect pixels are eliminated by thresholding. Figure 2 shows the iris pixel histograms for eyelash detection, the segmentation



Segmentation of Off-Axis Iris Images. **Figure 3** From [5], Fig. 3, p. 1170. ICE -1 file 244858, shown before and after the rotation process which corrects the iris to orthographic view. Copyright 2007 IEEE.

result, and the iris template after feature extraction (i.e., the IrisCode [2–4]).

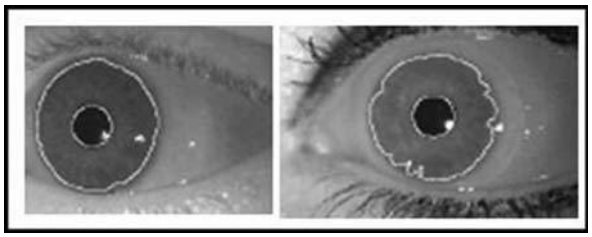
Alternative Segmentation Approaches

Some approaches to off-axis iris segmentation originate from research for ► eye tracking applications. There are many published eye tracking methods with various segmentation procedures, any of which could be borrowed or built upon for use in iris recognition. For present purposes, we mention the eye tracking methods of Zhu and Yang [6], and the related iris recognition paper by Sung et al. [7], which uses an initial step of eye corner detection (by filters which detect sideways V-shapes); then Sobel edge detection to find the nearly vertical left and right "sides" of the iris–sclera boundary, which also provides the center of the iris and approximate center of the pupil; intensity thresholding to identify dark regions; and finally, ► binary morphology (size and shape considerations) to choose the pupil disk from among the dark masses which lie within the previously determined iris boundary.

Zuo et al. [8] likewise use a combination of intensity and location for pupil segmentation, after a ► circular Hough transform which generates the list of pupil candidates. Additional morphology (finding a convex hull) eliminates indentations in the pupil mass caused by specular reflection at the edge of the pupil. Finally, the pupil boundary is declared to be an ellipse fitted to the resulting region. Fitting an ellipse to the pupil boundary is convenient for the eventual unwrapping of the iris, but the ellipse is used here also to find the limbic boundary: the operator (1) is modified so that, instead of families of circles, the contours of integration are ellipses of the same eccentricity and orientation as the pupil boundary.

Binary morphology may be the most ubiquitous idea in pupil segmentation; the structuring elements are easy to construct and combine in many convenient and intuitive ways. To cite a third example, Ross and Shah [9] find the pupil using thresholding to identify dark regions, median filtering to reduce noise, and a circle-fitting procedure to locate the circle containing the maximum number of black pixels, which is declared to be the pupil boundary. The main innovation in [9] is that the limbic boundary and eyelids are detected using geodesic active contours. The boundary is described as the zero level set of an embedding function, where the embedding function is initialized as the signed distance from any point in the image to a circle concentric with the pupil boundary, but of slightly larger radius. The active contour process makes the embedding function evolve according to gradients of the image grayscale values and the behavior of the stopping function. The resulting boundary contour traces the iris–sclera boundary arcs and the iris–eyelid boundary, creating the correct iris mask. Two sample images are shown in Fig. 4. Finally, the boundary is projected behind the eyelid by fitting a circle to the iris–sclera portion of the boundary contour, as this is needed for unwrapping.

In each of the segmentation methods mentioned so far, the segmentation process is performed prior to gaze angle determination: segmentation provides the information required to find the gaze angle using the eccentricity idea. However, the segmentation process and gaze angle determination can be intertwined in other ways. In each of the two methods proposed by



Segmentation of Off-Axis Iris Images. Figure 4 Iris segmentation result using geodesic active contours, from [9], Fig. 10a-b, shown on images from the West Virgina University off-angle iris image database [12]. Copyright 2006 IEEE.

Schuckers et al. [10], for instance, the rotation angle is incremented to create a series of possible image transforms. The angle is then selected according to an optimizing procedure. One of the ways to use this is by the application of camera calibration techniques to create model planes: plotting and tracking reference points from multiple eyes (image classes) at different angles, as a training process to transform other images. The other approach in [10] is to process the image through a projective transform incremented within a range of pitch and yaw angles. The rotation angle pair is then defined as the one which maximizes the operator (1) for the pupil, or in other words, the pitch and yaw combination which best transforms the pupil boundary back to a circle. This optimization process was also used by Dorairaj et al. in [11].

Summary

Segmentation is always the critical step for correct feature encoding. In the off-axis context, accurate segmentation is required not only to know which pixels to keep and which to throw away, but also (in most approaches) to ascertain the gaze angle, from which we can undo the effects of rotation before feature encoding takes place.

The most common themes and techniques that have been applied to on-axis segmentation appear just as often in off-axis segmentation, notably edge detection from large gradients, such as the Daugman operator (1), and various combinations of thresholding and binary morphology often brought to bear on finding the pupil: thresholding to identify dark regions, dilation and erosion for noise removal, convexity to remove specular reflection, and assorted ways to distinguish the pupil mass from the other dark regions by its size and near-roundness. The same issues that can plague on-axis segmentation are here as well, but it is to be expected that they will be exacerbated by the noncooperative image capture: blur, glare, eyelashes, and eyelids conspire to thwart off-axis segmentation and recognition.

There are no cut-and-dry "right answers" in iris segmentation; the iris–sclera boundary in particular is fairly subjective. Segmentation is often evaluated by visual inspection or after its incorporation into a complete recognition algorithm. In the latter case, the performance of segmentation is affected by the choice of the feature extraction process, the unwrapping process (assuming unwrapping is required), and in the off-axis case specifically, the results reflect the interactions between the segmentation, the unwrapping, and the determination of the gaze angle. Ultimately it is true that segmentation must perform within a start-to-finish recognition process, so all-up testing is a necessary step, but it is important to recognize the difficulty of cross-comparing segmentation performance statistics in isolation.

The handling of off-axis segmentation is a relatively new problem within the iris recognition community; the methods discussed in this chapter were mostly published in 2006–2007. For future work, a few obvious avenues of improvement are apparent. First, it has been an educational experience for the author of this chapter to explore a little more of the eye tracking world, since they have been hard at work on quick and robust segmentation at arbitrary angles for several years. One such paper [6] was mentioned in this chapter, but one could ask whether we have gleaned everything we can from this and other image processing applications. Second, there are very few off-axis image databases. Most of the research discussed was performed on the West Virginia University database [12], the NIST Iris Challenge Evaluation (ICE) database, or small sets of images collected locally by the researchers. Of course, the more databases that can be assembled, the better. Lastly, considerations about the three-dimensional shape of the iris have until now taken a back seat to other questions. This is certainly understandable, since these modeling issues would not affect the comparison of two images taken at the same angle, which is the traditional on-axis iris recognition assumption. Similarly, corneal refraction will affect

off-axis recognition performance, a preliminary study of which was undertaken in [13]. When it comes to the more challenging real-world iris recognition situations, even seemingly small improvements in iris modeling may prove very helpful.

Related Entries

- ▶ Iris Encoding and Recognition
- ▶ Iris Features and Anatomy
- ▶ Iris Recognition, Overview
- ▶ Iris Segmentation Using Active Contours

References

1. Bonney, B.L., Ives, R.W., Etter, D.M., Du, Y.: Iris pattern extraction using bit-planes analysis. In: Proceedings of the 38th Asilomar Conference on Signals, Systems and Computers, pp. 582–586 (Nov. 2004)
2. Daugman, J.G.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. **15**(11), 1148–1161 (1993)
3. Daugman, J.G.: Biometric personal identification system based on iris analysis. US Patent 5,291,560, US Patent Office, Washington, D.C., 1 Mar 1994
4. Daugman, J.G.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 21–30 (2004)
5. Daugman, J.: New methods in iris recognition. IEEE Trans Syst Man Cybern B **37**(5) (2007)
6. Jie Zhu, Jie Yang: Subpixel eye gaze tracking. In: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 124–129, Washington, D.C. (May 2002)
7. Eric Sung, Xilin Chen, Jie Zhu, Jie Yang: Towards non-cooperative iris recognition systems. In: Proceedings of the 2002 Seventh IEEE International Conference on Control, Automation, Robotics, and Vision, pp. 990–995, (Dec. 2002)
8. Jinya Zuo, Kalka, N.D., Schmid, N.A.: A robust iris segmentation procedure for unconstrained subject presentation. In: Special Session on Research at the Biometric Consortium Conference, 2006 Biometrics Symposium.
9. Ross, A., Shah, S.: Segmenting non-ideal irises using geodesic active contours. In: Special Session on Research at the Biometric Consortium Conference, 2006 Biometrics Symposium.
10. Schuckers, S.A.C., Schmid, N.A., Abhyankar, A., Dorairaj, V., Boyce, C.K., Hornak, L.A.: On techniques for angle compensation in nonideal iris recognition. In: IEEE Trans. Syst. Man Cybern. B **37**(5) (2007)
11. Dorairaj, V., Schmid, N.A., Fahmy, G.: Performance evaluation of non-ideal iris based recognition system implementing global ICA encoding. In: Proceedings of the IEEE International Conference on Image Processing, vol. 3, pp. 285–288 (2005)
12. Crihalmeanu, S., Ross, A., Hornak, L., Schuckers, S.: A Protocol for Multibiometric Data Acquisition Storage and Dissemination, Technical Report No. 5396, West Virginia University (2007)
13. Price, J.R., Gee, T.F., Paquit, V., Tobin, K.W. Jr.: On the efficacy of correcting for refractive effects in iris recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (June 2007)

Semi-Supervised

Part of data are labeled, but the rest are unlabeled. Algorithms are designed for separation of data underlying different classes, by taking the use of both labeled information and unlabeled information.

- ▶ Linear Dimension Reduction

Sensitivity Analysis in Biometric Systems

Sensitivity analysis in application to biometric systems investigates the influence of variations of certain parameters on the recognition performance of biometric systems. As an example, in the case of iris sample synthesis the selected parameters may include but are not limited to parameters characterizing fibers, collarette, and global features of iris images.

- ▶ Iris Sample Synthesis

Service-Oriented Architecture

Service-oriented architecture (SOA) are software architectures in which reusable services are deployed onto application servers and then consumed by clients in different applications or business processes [1]. SOA is a paradigm for organizing and utilizing distributed capabilities that may be under the control of different ownership domains. It provides a uniform means to offer, discover, interact with, and use capabilities to

produce desired effects consistent with measurable preconditions and expectations.

► Biometric Identity Assurance Services

Service Provider

In general, a business organization that provides consumers with specified services, for a fee or sometimes in response to a request. In the context of a federated registered traveler program, service providers contract with sponsoring entities (e.g., airports) to provide enrollment and/or verification services. These services generally include marketing, application processing, applicant registration (including data collection/biometric capture), data management/protection, card production and issuance/re-issuance, and overall program administration including fee collection. Verification services include provision of manned kiosks at travel lanes (e.g., screening checkpoints, passport control) with hardware/software to perform card reading/validation, biometric verification (capture/matching), security, and revocation checking.

► Registered Traveler

Session and Channel Variabilities

► Session Effects on Speaker Modeling

Session Effects on Speaker Modeling

Driss Matrouf, Jean-François Bonastre
CERI, University of Avignon, Avignon, France

Synonyms

Session and channel variabilities; Factor analysis; Nuisance attribute projection, Support vector machine; Score normalization

Definition

In spite of the research efforts in the fields of speaker feature characterization and speaker modeling, the speaker recognizer must face a problem involving the change of acoustic conditions, which vary in an unforeseeable way, from one recording to another. This phenomenon is generally referred to as *session mismatch* or *session variability*. It is one of the greatest sources of automatic speaker recognition (ASR) performance degradation. The term *session variability* encompasses a number of phenomena including transmission channel effects, environment noise (other people, cars, TV, etc.), differing room acoustics (hall, park, etc.), the microphone position relative to the mouth, and the variability introduced by the speaker himself. These sources of mismatch have the potential to increase the rate of errors of ASR. It is not easy to separately solve the problem of session variability caused by each of these sources thus, all the solutions suggested in literature deal with the problem of session mismatch in its globality. These proposed solutions perform at various levels of the ASR (feature space, model space and score space).

Introduction

In speaker recognition most of the errors are due not to the similarity among voiceprints of different speakers, but to the intrinsic variability of different utterances of the same speaker. This awkward variability for the speaker recognition is called *session variability*. One of the largest challenges in speaker recognition applications is dealing with session variability [1]. Session variability, also known as channel variability, is defined as the variability exhibited by given speaker from one recording session to another and is caused by several factors: channel effects, transducer characteristics, environment noise and intraspeaker variability. When this mismatch is produced between sessions extracted for enrolment and verification phases, performance of speaker recognition system is highly degraded and numerous techniques has been investigated to compensate session variability in several domains: In *feature domain*, compensation attempts at removing the channel effects from the feature vectors prior to modeling or testing. In *model domain*, compensation attempts to transform speaker models to minimize

the effects of varying channels. In *score domain*, compensation attempts to remove model score scales and shifts caused by varying input channel conditions.

Feature Domain Compensation

In literature concerning the variability of the session we find two types of compensation: unsupervised and supervised. In the unsupervised case, speaker features are transformed to accommodate the channel variation without a priori knowledge of the channel characteristics. On the other hand, supervised compensation estimates session-channel characteristics based on a priori knowledge of all possible channels.

Unsupervised Compensation: There are three types of unsupervised compensation. The first type use the temporal variability of feature vectors. It exploits the fact that the channel effects vary slowly with respect to the cepstral vectors (vocal tract characteristics evolution). For example, cepstral mean subtraction (CMS) [2, 3] subtracts the cepstral mean of an utterance from each of the cepstral vectors. ▶ RASTA-filtering [4] removes the slow varying components (corresponding to the channel) from the sequence of cepstral vectors by applying a bandpass filter. The second type of unsupervised compensation transforms the distorted features such that acoustic environments have minimum effect on the distribution of the transformed features. For example, in feature warping [6], observed features are mapped to a target distribution (e.g., standard normal) such that they follow the target distribution over a sliding window of feature vectors. In short-time Gaussianization [7], a linear transformation is applied to the distorted features before mapping them to a normal distribution. The transformation aims to decorrelate the feature vectors making them more suitable for diagonal covariance GMMs. It was found that short-time Gaussianization is superior to feature warping, especially at a low false acceptance rate. The third type makes use of the statistical difference between the clean acoustic models and the distorted speech to estimate a transformation matrix to map the distorted vectors to fit the clean model. This type of technique involves blind stochastic feature transformation [8].

Supervised Compensation: Supervised techniques are based on channel type detection during recognition. Examples in this category are feature mapping [9], spectral-magnitude matching [10], and stochastic feature transformation [11]. In feature mapping, the handset type of the testing utterance is identified by a handset detector. Feature vectors are then mapped to the channel-independent space based on the closest Gaussian in the channel-dependent GMM. In spectral-magnitude matching [10], a nonlinear polynomial mapper is trained to minimize the mean-squared spectral magnitude error between speech arising from electret and carbon-button handsets. This mapping is shown to be good at minimizing mismatches caused by phantom formants, bandwidth widening, and spectral flattening due to channel nonlinearity. Stochastic feature transformation (SFT) [11] aims to transform the distorted features to fit the clean speech models by selecting the most appropriate pre-computed transformation matrix. It has been shown that SFT can be extended to nonlinear feature transformation to overcome the nonlinear distortion [11].

Feature mapping [9] is a supervised technique that relies on a handset detector which contains the information of all possible channels that the users may use during verification. In feature mapping, the transformation is based on the top-1 Gaussian only. Specifically, let GMM $\Lambda^{CDi} = \{\pi_j^{CDi}, \mu_j^{CDi}, \Sigma_j^{CDi}\}_{j=1}^{M}$ be an M-mixture channel-dependent GMM for channel i and GMM $\Lambda = \{\pi_j, \mu_j, \Sigma_j\}_{j=1}^{M}$, be an M-mixture channel-independent root model. The mapping of a distorted vector y in the channel-independent space is given by

$$x = (y - \mu_k^{CDi}) \frac{\sigma_k^{CD}}{\sigma_k^{CDi}} + \mu_k^{CD} \qquad (1)$$

x is the corresponding frame in the channel-independent space. To account for the effect of other Gaussian components on the transformed features, the transformation should be based on a weighted average of all Gaussian components, which leads to probabilistic feature mapping (PFM).

Score Domain Compensation

Since the study of Li and Porter [40], various kinds of score normalization techniques have been proposed in the literature. Several of these are briefly described in the following section.

Score normalization techniques have mainly been derived from the study of Li and Porter [40]. In their paper, large variances had been observed from both

distributions of client scores (intraspeaker scores) and impostor scores (interspeaker scores) during speaker verification tests. As a result of these observations, the authors proposed solutions based on impostor score distribution normalization in order to reduce the overall score distribution variance (both client and impostor distributions) of the speaker verification system. The basic idea of the normalization technique is to center the impostor score distribution by applying on each score generated by the speaker verification system the following normalization. Let $L_\lambda(X)$ denote the score for speech signal X and speaker model λ. The normalized score $\tilde{L}_\lambda(X)$ is then given as follows:

$$\tilde{L}_\lambda(X) = \frac{L_\lambda(X) - \mu_\lambda}{\sigma_\lambda} \quad (2)$$

where μ_λ and σ_λ are the normalization parameters for speaker λ. Both of these parameters are estimated from a (pseudo) impostor score distribution. Different possibilities are available to compute the impostor score distribution: Znorm, Hnorm, Tnorm, HTnorm, Cnorm, Dnorm. All these normalization techniques are decsribed in [1]. This family of normalization techniques is the most commonly used in ASR. They are directly derived from Eq. 2 in which the scores are normalized by subtracting the mean and then dividing by the standard deviation. In the following the principle of the most popular is given:

Znorm The zero normalization (Znorm) technique is directly derived from the work done in [12]. It was quickly incorporated in most speaker verification systems in the middle of the nineties. In practice, a speaker model is tested against a set of speech signals produced by a group of impostors resulting in an impostor similarity score distribution. Speaker-dependent mean and variance-normalization parameters are estimated from this distribution and used to normalize similarity scores yielded by the speaker verification system when running using Eq. 16. One of the advantages of Znorm is that the estimation of the normalization parameters can be performed offline during speaker model training.

Hnorm By observing that, for telephone speech, most of the client speaker models respond differently according to the handset type used during testing data recording, Reynolds [13] proposed a variant of Znorm technique referred to as handset normalization (Hnorm) to deal with handset mismatch between the training and testing phases. Here, handset-dependent normalization parameters are estimated by testing each speaker model against handset-dependent speech signals produced by impostors. During testing, the type of handset relating to the incoming speech signal determines the set of parameters to use for score normalization.

Tnorm Still based on the estimate of mean and variance parameters to normalize impostor score distribution, test-normalization (Tnorm), proposed in [14], differs from Znorm by the use of impostor models instead of test speech signals. During testing, the incoming speech signal is classically compared with claimed speaker model as well as with a set of impostor models to estimate impostor score distribution and normalization parameters consecutively. If Znorm is considered as a speaker-dependent normalization technique, Tnorm can then be termed test-dependent. As the same test utterance is used during both testing and normalization parameter estimation, Tnorm avoids mismatch between test and normalization utterances which is a possible issue encountered when using Znorm. Conversely, Tnorm has to be performed online during testing.

Model Domain Compensation

Gaussian Mixture Models (GMM) used in a GMM-UBM framework is perhaps one of the most common configurations found in speaker verification systems [1]. For a several years now, new techniques that take session variability (or speaker intra-variability) into account have emerged providing a significant increase in system performance.

Among the approaches aimed at reducing the effect of session variability, feature mapping was often used alongside channel-labelled data with the assumption of a discrete channel space. The novelty brought by the factor analysis model is that it assumes the channel (or session) variability space to be continuous. In this model, the session variability effect is incorporated in the speaker model through session-dependent GMM mean supervectors offsets, constrained in a low dimensional subspace.

A speaker model can be decomposed into three different components: a speaker-session-independent component, a speaker dependent component and a session dependent component. A GMM mean supervector is defined as the concatenation of the GMM component means. Let D be the dimension of the

feature space and MD the dimension of a mean supervector where M is the number of Gaussian in the GMM. A speaker and session independent model is usually estimated in speaker verification to represent the inverse hypothesis: the UBM model. Let this model be parameterized by $\theta = \{\mathbf{m}, \mathbf{\Sigma}, \alpha\}$. In the following, (h, s) indicates the session h of the speaker s. The factor analysis model, in our case the eigenchannel MAP estimator, can be written as:

$$\mathbf{m}_{(h,s)} = \mathbf{m} + \mathbf{D}\mathbf{y}_s + \mathbf{U}\mathbf{x}_{(h,s)}, \tag{3}$$

where $\mathbf{m}_{(h,s)}$ is the session-speaker dependent supervector mean, $\mathbf{D}$ is $MD \times MD$ diagonal matrix, $\mathbf{y}_s$ the speaker vector (a MD vector), $\mathbf{U}$ is the session variability matrix of low rank R (a $MD \times R$ matrix) and $\mathbf{x}_{(h,s)}$ are the channel factors, a R vector (theoretically $\mathbf{x}_{(h,s)}$ does not dependent on s). Both $\mathbf{y}_s$ and $\mathbf{x}_{(h,s)}$ are normally distributed among $\mathcal{N}(0, I)$. $\mathbf{D}$ satisfies the following equation $\mathbf{I} = \tau \mathbf{D}^t \mathbf{\Sigma}^{-1} \mathbf{D}$ where τ is the *relevance factor* required in the standard MAP adaptation ($\mathbf{D}\mathbf{D}^t$ represents the *a priori* covariance matrix of $\mathbf{y}_s$).

The success of the factor analysis model relies on a good estimation of the $\mathbf{U}$ matrix. This requires a sufficiently high amount of data in which a high number of different recordings per speaker are available.

The verification task is defined as follows. A speaker $\mathbf{s}_{tar}$ is enrolled by the system with his training data $Y_{\mathbf{s}_{tar}}$. Given a sequence of speech frames $\mathcal{Y} = \{y_1 \dots y_T\}$ and the speaker $\mathbf{s}_{tar}$, the speaker verification task consists of determining if $\mathcal{Y}$ was spoken by $\mathbf{s}_{tar}$ or not. Using the factor analysis decomposition in both training and testing data, one can write:

$$\begin{aligned} \mathbf{m}_{(\mathbf{h}_{tar}, \mathbf{s}_{tar})} &= \mathbf{m} + \mathbf{D}\mathbf{y}_{\mathbf{s}_{tar}} + \mathbf{U}\mathbf{x}_{\mathbf{h}_{tar}}, \\ \mathbf{m}_{(\mathbf{h}_{test}, \mathbf{s}_{test})} &= \mathbf{m} + \mathbf{D}\mathbf{y}_{\mathbf{s}_{test}} + \mathbf{U}\mathbf{x}_{\mathbf{h}_{test}}, \end{aligned} \tag{4}$$

where the speaker $\mathbf{s}_{tar}$ in the training data and $\mathbf{s}_{test}$ in the testing data have been distinguished. To address session variability, the strategy adopted by [15, 16] assumes that the test speaker has the same identity as the target speaker, *i.e.* $\mathbf{y}_{\mathbf{s}_{test}} = \mathbf{y}_{\mathbf{s}_{tar}}$. The channel component $\mathbf{U}\mathbf{x}_{\mathbf{h}_{test}}$ of the test segment is estimated under this assumption. Indeed, the session component of the target model $\mathbf{U}\mathbf{x}_{\mathbf{h}_{tar}}$ is replaced by the one estimated from the test data $\mathbf{U}\mathbf{x}_{\mathbf{h}_{test}}$. The world model in the score equation remains unchanged. This strategy has several drawbacks: the target speaker model is changed for each test and significant performance gains can only be achieved when score normalization techniques are employed. In [17], a hybrid domain normalization strategy is proposed which aims to withdraw the session component in the test and training data. This can be formulated as,

$$\mathbf{m}_{\mathbf{s}_{tar}} = \mathbf{m} + \mathbf{D}\mathbf{y}_{\mathbf{s}_{tar}}; \ \mathbf{m}_{\mathbf{s}_{test}} = \mathbf{m} + \mathbf{D}\mathbf{y}_{\mathbf{s}_{test}}. \tag{5}$$

In this strategy, speakers are assumed to be different and are treated separately. When using a LLR-based verification approach, the speaker verification score is an expected log-likelihood ratio:

$$LLK(\mathcal{Y}|\mathbf{m}_{(\mathbf{h}_{tar}, \mathbf{s}_{tar})}) - LLK(\mathcal{Y}|\mathbf{m}), \tag{6}$$

where $LLK(\cdot|\cdot)$ indicates the average of the log-likelihood function over all frames. Here, all GMMs have the same covariance matrices as well as the same mixture weights (both dropped from the equation for clarity). Two session compensation approaches can be adopted. The first approach involves performing compensation at the frame level, where session compensation can be seen as a front-end process. The second approach is a hybrid compensation, where the session variability is subtracted from the target speaker model (model domain) and the compensation in the testing data is performed at the frame level (feature domain). The following formula is used to remove the session effect for each frame t (also successfully used by [18]):

$$\hat{t} = t - \sum_{g=1}^{M} \gamma_g(t) \cdot \{\mathbf{U} \cdot \mathbf{x}_{\mathbf{h}_{test}}\}_{[g]}. \tag{7}$$

A similar approach, NAP (nuisance attribute projection) [5] is also used to reduce the impact of channel (or session) [5]. NAP is dedicated to the ▶ SVM supervector (support vector machine) based systems [22], which called the ▶ NAP-SVM approach.

Experiments

Speaker verification experiments are performed using the NIST SRE 2005 as a development set and the 2006 database for the validation set using only male speakers. These databases consist entirely of telephony speech and are referred to as the 2005 and 2006 protocols. The 2005 protocol consists of 274 speakers, 9,012 tests (951 target tests, the remainder being impostor trials) while the 2006 protocol consists of 354 speakers, 9,720 tests (of which 741 were target tests).

(Please note that 2005 protocol corresponds to the core condition labeled as *det7* and the 2006 protocol corresponds to the core condition labeled as *det3*.) Results are given in terms of equal-error-rate (EER) and the minimum DCF (an *a posteriori* decision). On average, train and test utterances contain 2.5 min of speech where around 30% of speech frames per speaker have been retained. The intersession variability matrix is trained on the NIST-SRE-2004 database with 2,938 examples from 124 speakers, taking approximately 20 iterations to reach convergence.

The baseline system is a standard GMM-UBM system [19]. Frames are composed of 19 LFCC parameters, their derivatives, and 11 second derivatives (the frequency window is restricted to 300–3,400 Hz). A normalization process is applied so that the distribution of each cepstral coefficient is 0-mean and 1-variance for a given utterance. Table 1 shows the results of the baseline system and T-norm scores (Z and ZT-norm do not bring any improvement).

The following outlines the experimental results obtained with the implementation described in [17]. Table 2 shows the improvement with score normalization. Znorm consistently brings about an improvement in both protocols, while Tnorm is only effective in 2006. Indeed, the DCFmin drops from 1.83 to 1.64 with Znorm in 2005 and from 1.61 to 1.18 with ZTnorm in 2006. While the behavior is different on both protocols, ZTnorm appears to be the most confident choice for score normalization.

Table 1 shows the results of the baseline system and T-norm scores (Z and ZT-norm do not bring any improvement).

Session Effects on Speaker Modeling. **Table 1** Results of the baseline GMM-UBM system on the 2005 and 2006 protocol. DCFmin (× 100), EER(%)

	SRE-05		SRE-06	
	DCFmin	EER	DCFmin	EER
Nonorm	3.83	7.15	3.88	6.79
Tnorm	3.05	8.52	2.9	5.7

Session Effects on Speaker Modeling. **Table 2** Baseline and score normalized results obtained when using the factor analysis model (rank = 40). Znorm consistently provides an improvement over the baseline results. DCFmin (× 100), EER(%)

	SRE-05		SRE-06	
	DCFmin	EER	DCFmin	EER
Nonorm	1.83	4.42	1.61	2.97
Tnorm	1.84	4.72	1.29	2.83
ZTnorm	1.72	4.62	1.18	2.15
Znorm	1.64	4.21	1.46	2.33

Conclusion

The majority of techniques presented in this chapter fail to reach their goal for different reasons. For example, feature mapping, speaker model synthesis (SMS) [20], and H-Norm are sub-optimal because they consider only finite and discrete sources of session mismatch. In theory, assuming that the mismatch belongs to a finite number of categories simplifies dealing with the session variability problem. However, when the session characteristics are very different from those considered, inappropriate normalisation is applied which can cause recognition or verification error. When using this technique it is necessary to automatically detect the appropriate category for each test utterance. This process is somewhat error-prone having the potential to cause verification errors when the inappropriate normalisation in applied.

The second key deficiency is not actually modeling the effects of session variability but simply attempting to suppress them. Feature warping, T-Norm and Z-Norm fit into this category. These methods have no knowledge of the specific conditions encountered in a recording but use some a-priori information of the effects caused by session variability.

For several years, new approachs have emerged that take session variability (or speaker intra-variability) into account, providing a significant increase in system performance. The main feature of these techniques is the way in which the issue of mismatch in GMM-based speaker verification in addressed. This is achieved through the explicit modelling of session variability in both the training and testing procedures. These directly model the mismatch between sessions in a constrained subspace of the GMM speaker model means. This replaces the discrete categorisation of techniques such as feature mapping and H-Norm with a continuous

vector-valued representation of the session conditions. A key strength of this approach is the avoidance of data labelling requirements due to the particular training methods that are employed.

Related Entries

- ▶ Gaussian Mixture Models
- ▶ Speaker Matching
- ▶ Speaker Recognition, Overview
- ▶ Uiniversal Background Models

References

1. Bimbot, F., Bonastre, J.-F., Fredouille, C., Gravier, G., Magrin-Chagnolleau, I., Meignier, S., Merlin, T., Ortega-Garcia, J., Petrovska, D., and Reynolds, D.A.: A tutorial on text-independent speaker verification. EURASIP J. Appl. Signal Process., Special issue on biometric signal processing (2004)
2. Atal, B.S.: Effectiveness of LPC characteristics of the speech waves for A.S.I and A.S.V. J. Acoust. Soc. Am. JASA **55** (1974)
3. Furui, S.: An analysis of long-term variation of feature parameters of speech and its application to talker recognition. Electron. Commun. **57-A**, 34–42 (1977)
4. Hermansky, H., Morgan, N.: RASTA processing of speech. IEEE Trans. Speech Audio Process. **2**, 578–589 (1994)
5. Solomonoff, A., Campbell, W.M., and Boardman, I.: Advances in channel compensation for SVM speaker verification. In Proceedings of ICASSP (2005)
6. Pelecanos, J., Sridharan, S.: Feature warping for robust speaker verification. 2001: a Speaker Odyssey. The Speaker Recognition Workshop, Chania, Greece, pp. 213–218. Chania, Crete (2001)
7. Xaing, B., Chaudhari, U., Navratil, J., Ramaswamy, G., and Gopinath, R.: Short-time Gaussianization for robust speaker verification. In: Proceedings of International Conference on Acoustics Speech and Signal Processing (ICASSP), Vol. 1, pp. 681–684 (2002)
8. Yiu, M.C.K.K., Mak, M.W., Kung, S.: Blind stochastic feature transformation for channel robust speaker verification. J. VLSI Signal Process. **42**(2), 117–126 (2006)
9. Reynolds, D.: Channel robust speaker verification via feature mapping. In: Proceedings of International Conference on Acoustics Speech and Signal Processing (ICASSP 2003), vol. 2. Hong Kong, China (2003)
10. Quatieri, D.A.R.T.F., OLeary, G.C.: Estimation of handset nonlinearity with application to speaker recognition. IEEE Trans. Speech Audio Process. **8**(5), 567–584 (2000)
11. Kung, S.Y., Mak, M.W., and Lin, S.H.: Biometric Authentication: A Machine Learning Approach. New Jersey: Prentice-Hall, Upper Saddle River (2005)
12. Li, K.P., Porter, J.E.: Normalizations and selection of speech segments for speaker recognition scoring. In: Proceedings of International Conference on Acoustics Speech and Signal Processing (ICASSP 98), pp. 595–598 (1988)
13. Reynolds, D.A.: The effects of handset variability on speaker recognition performance: experiments on the switchboard corpus. In: International Conference on Acoustics, Speech, and Signal Processing ICASSP (1996)
14. Auckenthaler, R., Carey, M., Lloyd-Thomas, H.: Score normalization for text-independent speaker verification system. Digital Signal Processing (DSP), a review journal – Special issue on NIST 1999 speaker recognition workshop **10**(1–3), 42–54 (2000)
15. Kenny, P., Boulianne, G., Ouellet, P., Dumouchel, P.: Factor Analysis Simplified. In: Proceedings of International Conference on Acoustics Speech and Signal Processing (ICASSP 2005), vol. 1. Philadelphia, USA (2005)
16. Vogt, R., Baker, B., Sridharan, S.: Modelling Session Variability in Text-Independent Speaker Verification. In: Proceedings of Interspeech, European Conference on Speech Communication and Technology (Eurospeech 2005). Lisboa, Portugal (2005)
17. Matrouf, D., Scheffer, N., Fauve, B., Bonastre, J.-F.: A straightforward and efficient implementation of the factor analysis model for speaker verification. In INTERSPEECH Conference. Antewerp, Belgium (2007)
18. Vair, C., Colibro, D., Laface, P.P.: Channel factors compensation in model and feature domain for speaker recognition. In: Odyssey'06, the Speaker Recognition Workshop. San Juan, Puerto Rico (2006)
19. Reynolds, D.A., Quatieri, T.F., Dunn, R.B.: Speaker verification using adapted gaussian mixture models. Digit. Signal Process. (DSP), a review journal – Special issue on NIST 1999 speaker recognition workshop, **10**(1–3), 19–41 (2000)
20. Teunen, R., Shahshahani, B., and Heck, L.: A model-based transformational approach to robust speaker recognition. In: International Conference on Spoken Language Processing, pp. 495–498 (2000)
21. Hanson, B.A. et al.: Subband or cepstral domain filtering for recognition of Lombard and channel distorded speech. In Proceedings of ICASSP-93, pp. 79–82 (1993)
22. Collobert, R., Bengio, S.: SVM Torch: support Vector Machines for large-scale regression problems. J. Mach. Learn. Res. **1**, 143–160 (2001)

SFinGe

Raffaele Cappelli
Biometric System Laboratory – DEIS – University of Bologna

Synonyms

Synthetic Fingerprint Generator

Definition

SFinGe (Synthetic Fingerprint Generator) is a fingerprint sample synthesis approach developed by the Biometric System Laboratory of the University of Bologna (Italy). It is available as a software program able to generate large databases of images very similar to human's fingerprints, together with ► ground-truth data about their characteristics and features. These databases are particularly useful for developing, optimizing and testing fingerprint recognition systems and are being extensively used by industrial, academic and government organizations.

Overview and History

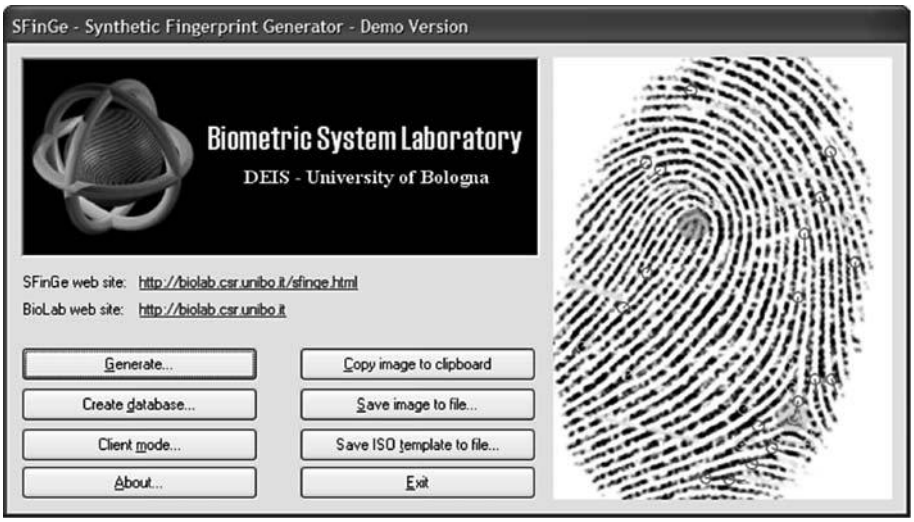
SFinGe (the Italian for *sphinx*, pronunciation *sphin-je*) is the acronym for Synthetic Fingerprint Generator. SFinGe can be used to easily create large databases of fingerprints, thus allowing recognition algorithms to be simply trained, tested and optimized. The images generated emulate fingerprints acquired with on-line sensors (see ► Fingerprint Acquisition) but, with a few changes, the simulation of impressions produced by the ink-technique is also possible.

This ► fingerprint sample synthesis approach was developed at the Biometric System Laboratory of the University of Bologna (Italy) [1] since 1999; the first version of the method [2] was able to synthesize realistic fingerprint patterns, but limited to only one impression of each "synthetic finger". Appropriate techniques for simulating more impressions of the same finger were developed at the end of 2000, successfully adopted to generate one of the test databases for the first Fingerprint Verification Competition [3], and described, for the first time, in [4]. In 2002, realistic background generation capabilities were added [5] and, in 2004, an improved noise model was developed [6]. More recently, the approach has been expanded with the generation, for each synthetic fingerprint image, of ► ground-truth minutiae information (i.e., the precise location and characteristics of each minutia) and other features (such as the ► orientation field) [7].

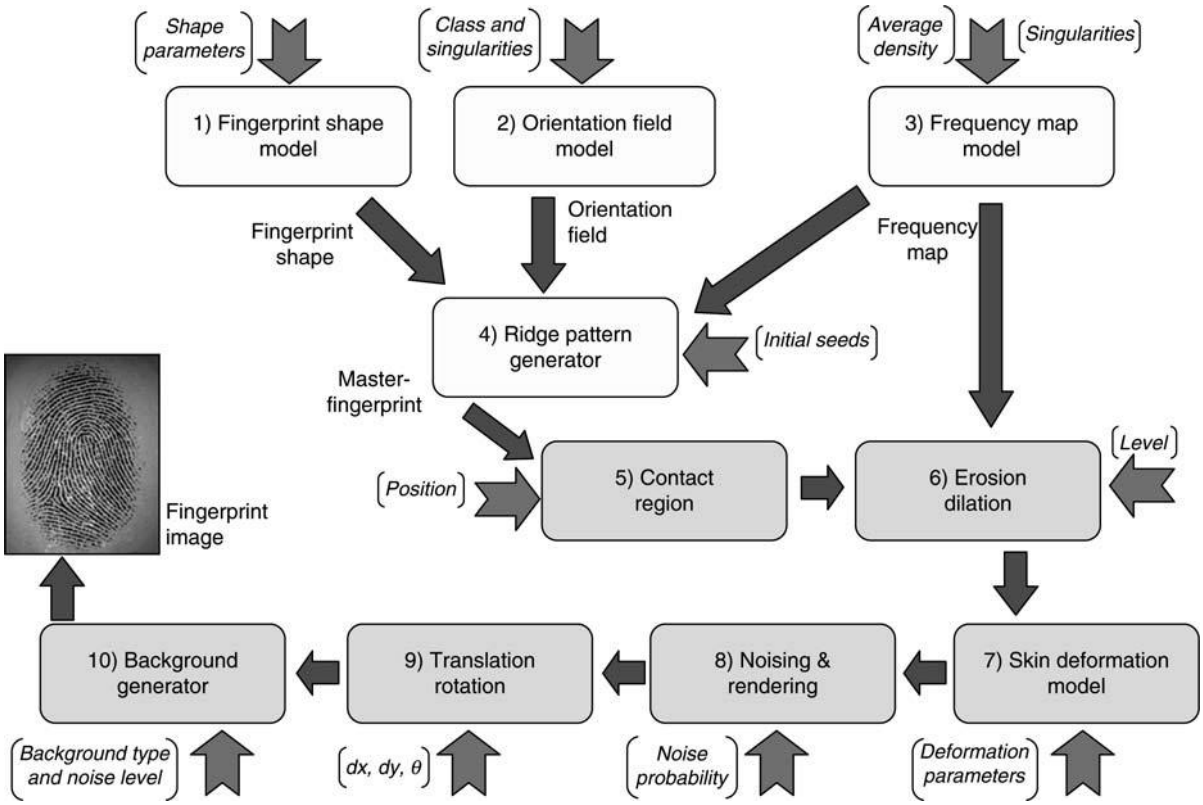
A software tool for generating fingerprint images according to the SFinGe method has been provided by the Biometric System Laboratory of the University of Bologna since 2001. A demo version of this tool (Fig. 1) can be downloaded from http://biolab.csr.unibo.it/sfinger.html and has been used to prepare most of the figures in the following sections.

The Generation Process

Fig. 2 shows a functional schema of the generation approach. SFinGe adopts a statistical ridge pattern model (see fingerprint sample synthesis) to create a *master-fingerprint* (that is the unique and immutable



SFinGe. Figure 1 The user interface of the SFinGe software tool.



SFinGe. **Figure 2** A functional schema of the SFinGe generation approach: each rounded-box represents a step (based on a corresponding mathematical model); the main parameters are reported between square brackets. Steps 1–4 create a master-fingerprint, steps 5–10 generate the final synthetic image.

characteristics of a "synthetic finger") through the following steps [5]:

1. *Fingerprint shape generation*: Definition of the global shape of the fingerprint, according to a simple model based on elliptical segments;
2. *Orientation field generation*: A mathematical ridge-flow model allows to generate a consistent orientation field;
3. *Frequency map generation*: The local ridge-line frequency is generated on the basis of some heuristic criteria;
4. *Ridge-line pattern generation*: Ridge-lines and minutiae are created using space-variant filtering.

Once a master-fingerprint has been created, one ore more of its "impressions" can be randomly generated, by applying the following steps [5]:

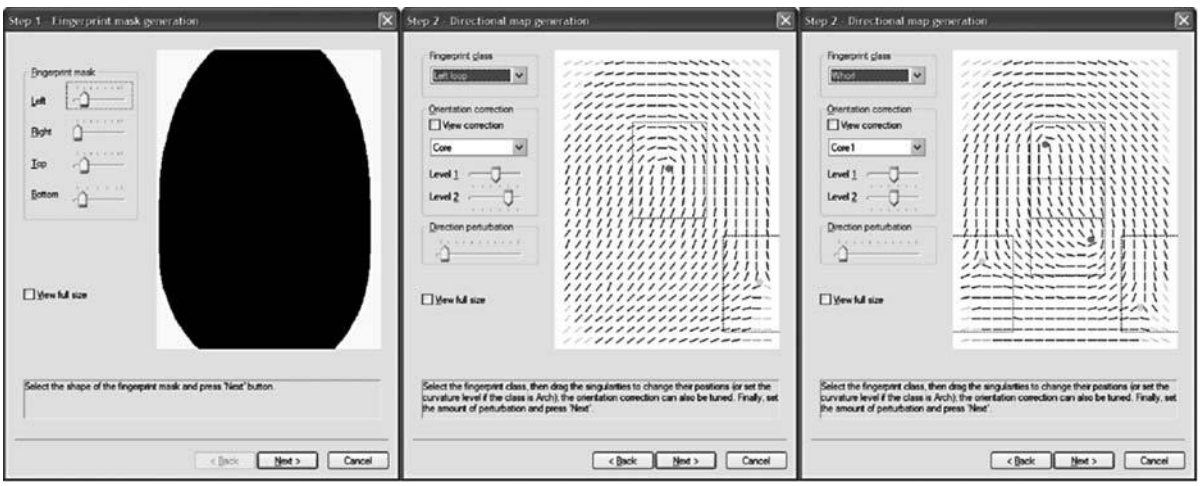
5. *Selection of the contact region*: The ridge-line pattern is translated without modifying the global fingerprint shape and position (this simulates different finger placements over the acquisition device);
6. *Variation of the ridge-line thickness*: Morphological operators are applied to simulated different degrees of skin dampness and/or finger pressure;
7. *Fingerprint distortion*: A skin distortion model is adopted to simulate the effects of skin elasticity;
8. *Noising and rendering*: A gray-scale noisy image is produced by modeling some of the factors that deteriorate the quality of real fingerprints;
9. *Global translation/rotation*: The image is randomly translated and/or rotated, to simulate real fingerprints that usually are not perfectly centered and can present a certain amount of rotation;
10. *Background generation*: A realistic background can be created to simulate a given acquisition device.

Fig. 2 shows, for each of the steps described above, the various input parameters (graphically represented by the red arrows). The SFinGe software tool lets the user

adjust most of those parameters and observe the corresponding effects on the synthetic fingerprint (see Figs. 3–7). The software also allows a database of synthetic fingerprints to be batch-generated, given a relatively small set of input parameters (see Fig. 8), including: number of fingers, impressions per finger, image size, seed for the random number generator, maximum amount of noise, maximum amount of deformation. During the batch-generation of a fingerprint database, each master-finger is generated by using a different seed for the random number generator; those seeds are randomly selected as well. During the generation of a single database, all the seeds chosen are different; although it is reasonable to assume that different seeds imply different fingerprints, it might happen that two different seeds produce almost identical fingerprint images. To reduce this hypothetical risk, SFinGe adopts one of the best ▶ pseudo-random number generators proposed in the scientific literature [8].

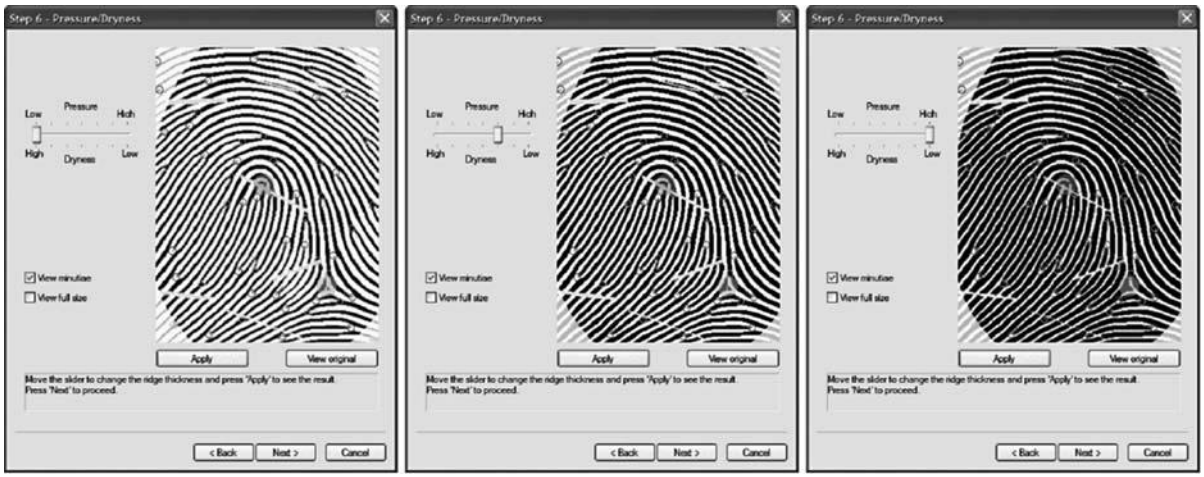
The creation of minutiae ground-truth proceeds in parallel with the fingerprint generation (Fig. 9): the standard minutiae extraction procedure defined in [9] is applied to the master-fingerprint, then all the relevant transformations executed on the fingerprint are applied to the minutiae (e.g., translation, rotation, distortion). This approach has some clear advantages:

1. The features can be extracted by applying the standard procedures easily and without ambiguities, since the extraction occurs on a binary image without any noise;
2. The ground truth is always unique and sound, even when the quality of the final image is very low (see Fig. 10).



SFinGe. **Figure 3** Graphical user interface for the fingerprint shape generation (left) and orientation field generation (middle and right).

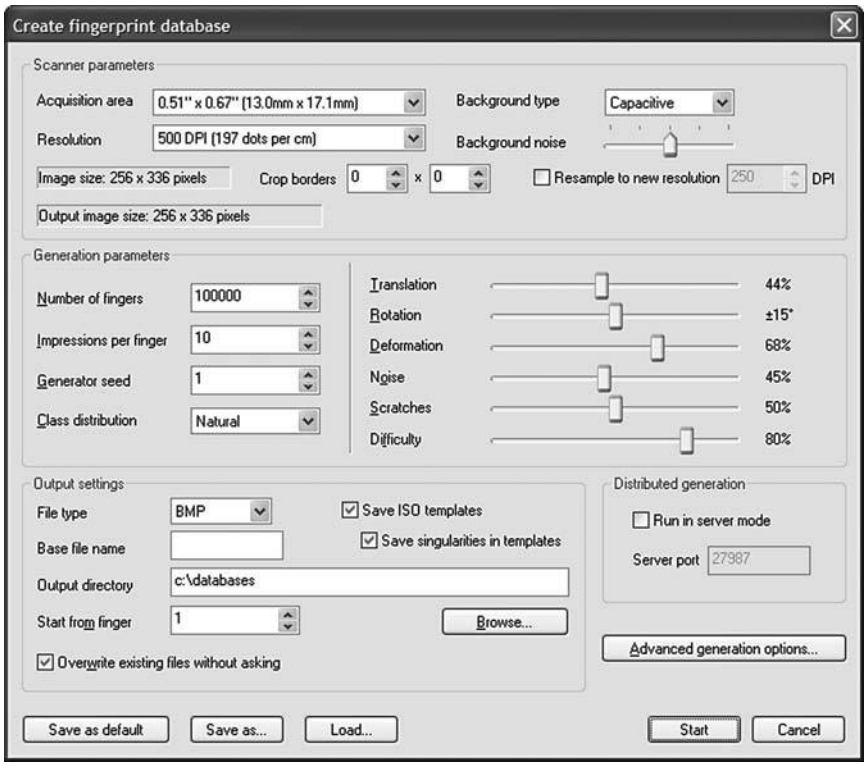
SFinGe. **Figure 4** Graphical user interface for the ridge-line pattern generation.



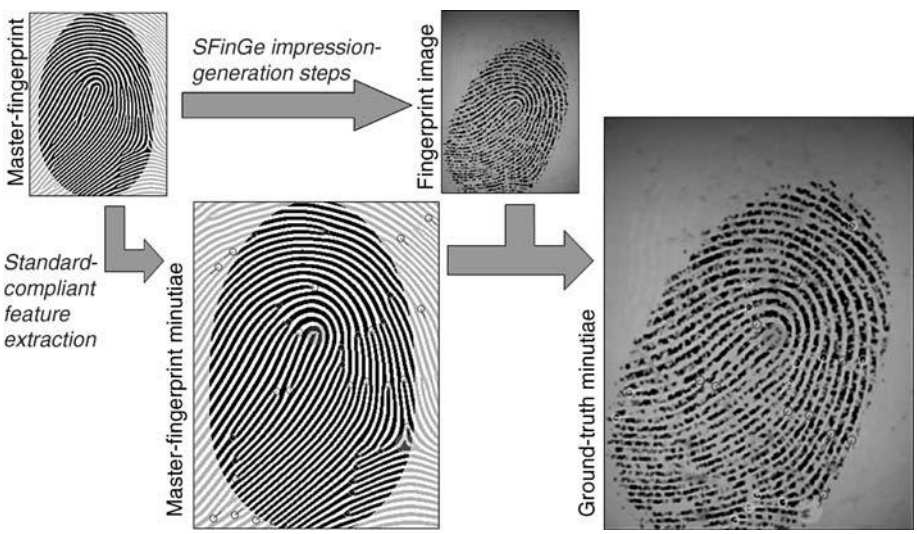
SFinGe. **Figure 5** Graphical user interface for the variation of the ridge-line thickness.

SFinGe. **Figure 6** Graphical user interface for fingerprint distortion (left) and noising (middle and right).

SFinGe. **Figure 7** Graphical user interface for global rotation/translation (left) and background generation (middle and right).



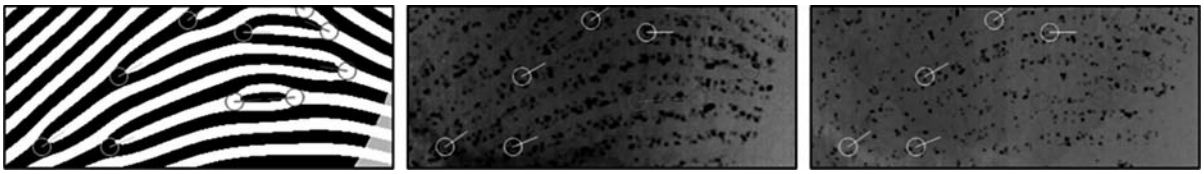
SFinGe. **Figure 8** The batch-generation options.



SFinGe. **Figure 9** Generation of ground-truth minutiae data.

Generation of other ground-truth features is performed in a similar fashion; for instance, all the relevant transformations can be applied to the orientation field calculated at step 2, thus obtaining the true orientation field of the final synthetic fingerprint.

The automatic generation of a whole fingerprint database (including ground-truth data) is totally parallelizable, since the generation of each master-fingerprint (with its impressions) is independent of the others. This makes it possible to distribute the



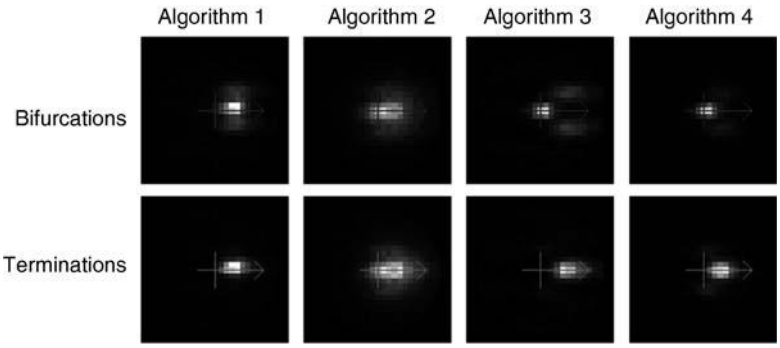
SFinGe. **Figure 10** Minutiae ground-truth as generated by SFinGe for very high-quality fingerprints (left), medium-quality fingerprints (middle), and low-quality fingerprints (right).

process on many computers; for instance, using ten 3GHz PCs in a network, a database of 100,000 fingerprints (10,000 fingers, 10 impressions per finger) can be generated in less than two hours.

Applications

SFinGe can be used to create, at zero cost at without any ► privacy issue, large databases of fingerprints, whose main characteristics can be controlled and adjusted according to the specific needs of a given application. Furthermore, ground-truth data about the main fingerprint features can be automatically produced for each fingerprint in the database. SFinGe has been used by many industrial, academic and government organizations; the main applications of this synthesis approach are described in the following.

1. *Performance evaluation*: SFinGe is an effective tool to overcome the problem of collecting large fingerprint databases for test purposes. Obviously real fingerprint databases cannot be completely substituted, especially when performance has to be measured referring to a given real environment/application; on the other hand synthetic fingerprints proved to be well suited for ► technology evaluations [10]: a comparison of the behavior of several fingerprint matching algorithms on real and synthetic databases showed that not only the performance is very similar, but the genuine/impostor distributions and the FMR/FNMR curves (see ► Performance Measures) are also surprisingly close [5].
2. *Training*: Many classifiers and pattern recognition techniques (e.g., neural networks, ► Principal Component Analysis, ► Support Vector Machines) require a large training set for an accurate learning stage. Synthetic fingerprint images are very well suited to this purpose: in fact the generator input parameters allow to explicitly control the type and features of the generated sets (e.g., class, type of noise, distortion) and this can be exploited in conjunction with boosting techniques to drive the learning process. For example, in [11], a large synthetic training set (generated by SFinGe) was successfully used to derive optimal parameters for ► fingerprint indexing.
3. *Security evaluation*: Synthetic fingerprints can be used to test the robustness of fingerprint verification systems to "Trojan horse" attacks against the sensor or the feature extractor [5] (see ► Biometric Security, Overview). SFinGe allows to generate large sets of fingerprints whose features (e.g. minutiae distribution) can be varied independently of other fingerprint characteristics (e.g. orientation field), and therefore, it is well suited for studying the robustness against "hill-climbing" attacks (see [5]).
4. *Semantic conformance to standards*: Interoperability tests [12, 13] have shown that the location, direction, and type of minutiae extracted by different minutiae extraction algorithms from the same finger image tend to be different (see ► Finger Data Interchange Format, Standardization). Algorithms syntactically compliant to standards such as the ISO/IEC 19794-2 [9], are often not semantically compliant and this creates huge interoperability problems. Unfortunately, testing semantic conformance to a minutiae extraction standard is not easy, since it requires a lot of data with manually-labeled minutiae points (ground-truth); furthermore, in low-quality areas, even the manual labeling of minutiae points is not reliable. The automatic generation of ground-truth data for synthetic fingerprint images provided by SFinGe is an effective way to carry out semantic conformance and interoperability studies. For instance, in [7] a synthetic database has been used to analyze the distribution of minutiae positions and directions of some algorithms with respect to the ground-truth (see Fig. 11).



SFinGe. **Figure 11** Distributions of minutiae placement and direction as estimated in [7] for some feature extraction algorithms. In each image, the intensity I[*x*, *y*] is proportional to the estimated likelihood that a minutia will be found by an algorithm at position (*x*, *y*) with respect to the ground-truth minutia direction (denoted by the arrow).

Related Entries

- ▶ Anatomy of Fingerprint
- ▶ Finger Data Interchange Format, Standardization
- ▶ Fingerprint Classification
- ▶ Fingerprint Databases and Evaluation
- ▶ Fingerprint Features
- ▶ Fingerprint Indexing
- ▶ Fingerprint Noise
- ▶ Fingerprint Orientation Field
- ▶ Fingerprint Ridge-line Pattern
- ▶ Fingerprint Sample Synthesis
- ▶ Fingerprint Singularities, Minutiae, Pores
- ▶ Performance Measures
- ▶ Privacy Issues
- ▶ Support Vector Machine

References

1. Biometric System Laboratory (University of Bologna) web site, Available at: http://biolab.csr.unibo.it. Accessed January 2008
2. Cappelli, R., Erol, A., Maio, D., Maltoni, D: "Synthetic Fingerprint-image Generation." In: Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, vol. 3, pp. 475–478 September (2000)
3. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2000: Fingerprint Verification Competition. IEEE Trans Pattern Anal Mach Intell **24**(3), 402–412 (2002)
4. Cappelli, R., Maio, D., Maltoni, D: "Synthetic Fingerprint-Database Generation." In: Proceedings of the 16th International Conference on Pattern Recognition, Québec City, vol. 3, pp. 744–747, August (2002)
5. Cappelli, R: Synthetic fingerprint generation, In: Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S. (eds.) Handbook of Fingerprint Recognition, Springer, New York, (2003)
6. Cappelli, R., Maio, D., Maltoni, D., "An Improved Noise Model for the Generation of Synthetic Fingerprints." In: Proceedings Eighth International Conference on Control, Automation, Robotics and Vision (ICARCV2004), Kunming, China, December (2004)
7. Cappelli, R: "Use of Synthetic Data for Evaluating the Quality of Minutia Extraction Algorithms", In: Proceedings of the Second NIST Biometric Quality Workshop, Gaithersburg, Maryland, November (2007)
8. Saito, M., Matsumoto, M: "SIMD-oriented Fast Mersenne Twister: a 128-bit Pseudorandom Number Generator", Monte Carlo and Quasi-Monte Carlo Methods 2006, pp. 607–622. Springer, New York, (2008)
9. ISO/IEC 19794-2:2005, Information technology – Biometric data interchange formats – Part 2: Finger minutiae data
10. Cappelli, R., Maio, D., Maltoni, D., Wayman, J.L., Jain, A.K. Performance Evaluation of Fingerprint Verification Systems. IEEE Trans Pattern Anal Mach Intell. **28**(1), 3–18 (2006)
11. Cappelli, R., Maio, D., Maltoni, D: "Indexing Fingerprint Databases for Efficient 1:N Matching." In: Proceedings of the Sixth International Conference on Control, Automation, Robotics and Vision (ICARCV2000), Singapore, December (2000)
12. Minutiae Interoperability Exchange Test (MINEX) web site, Available at: http://fingerprint.nist.gov/minex04. Accessed January 2008
13. Minutiae Template Interoperability Testing (MTIT) Project web site, Available at: http://www.mtitproject.com. Accessed January 2008

Shape

"Shape is all the geometric information that remains when location, scale and rotational effects are filtered out from the object". Kendall's statistical shape is a sparse descriptor of the shape that describes

the shape configuration of k landmark points in an m-dimensional space as a $k \times m$ matrix containing the coordinates of the landmarks from which scale and translation are filtered out. This shape feature therefore lies on a spherical manifold which is well studied and therefore properly defined distance measures are available in this manifold to perform recognition.

▶ Gait Biometrics, Overview
▶ Gait Recognition, Model-Based

Shape Index

Shape index *Si*, a quantitative measure of the shape of a surface at a point p, is defined by (1),

$$S_i(p) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \qquad (1)$$

where $k1$ and $k2$ are maximum and minimum principal curvatures, respectively. With this definition, all shapes can be mapped into the interval $S_i \in [0,1]$. The larger shape index values represent convex surfaces and smaller shape index values represent concave surfaces.

▶ Ear Biometrics, 3D

Shape Model

Shape model is often embedded within a structural model. It uses geometrical shapes e.g., stick figures as a reduced representation of human body, blob, or cylinder to represent body masses, or silhouette outline of a human figure obtained from edge information to describe the body of interest.

▶ Gait Recognition, Model-Based

Shape vs. Texture

▶ Face Recognition, Geometric vs. Appearance-Based

Shielding Functions

A theoretical approach to cancelable biometrics developed by Linnartz et al. (2003). Shielding functions allow a verifier to check the authenticity of a prover (user wanting to be verified) without learning any biometric information. The scheme depends on proposed δ-contracting and ε-revealing functions, which allow testing whether measured features are within a range. Under some assumptions, a biometric may be tested without learning anything about the biometric feature values. One limitation of this scheme is that biometric samples are assumed to be perfectly registered. This scheme offers an interesting cryptographic basis for the construction of encrypted biometric systems, and it is used by several authors.

▶ Cancelable Biometrics

Shoeprint Matching

▶ Footwear Recognition

Shot Noise

For image sensors, photon conversion into electrons at a pixel is a random event. Random fluctuations in the number of electrons for a fixed number of photons are referred to as shot noise. The standard model for shot noise is that the noise is proportional to the square root of the number of electrons generated. Hence, the signal/noise ratio is proportional to $\sqrt{N}$, where N is the number of electrons.

▶ Biometric Sensor and Device, Overview
▶ Iris Device

Side-Channel Attacks

Side-channel attack is an attack against a cryptographic or biometric security system based on measurements

of the implementation of the system, rather than on weaknesses in the algorithms. For example, side-channel attacks may use timing, power consumption, or electromagnetic measurements on the security device. Side-channel attacks are primarily of concern for biometric encryption systems and match-on-card devices where an attack could potentially be mounted by iteratively improving the presented biometric. Very little research has been done to explore the feasibility of side-channel attacks, but the success of attacks on biometric template security and biometric encryption suggests that such attacks are certainly feasible.

► Security and Liveness, Overview

Signal to Noise Ratio

Information is transmitted or recorded by variations in a physical quantity. For any information storage or transmission system, there will be intended variations in the physical quantity – signal – and unintended variations – noise. In an analog telephone system, the signal (voice) is represented by variation in a voltage level. As the signal is transmitted along a phone line, it can pickup other unintended variations – e.g., leakage of other signals, static from electrical storms – that are noise. The ratio of the signal level to the noise level is the signal to noise ratio (SNR). Since it is a ratio, SNR is dimensionless. However, SNR can be expressed as either an amplitude ratio (voltage ratio for the phone example) or a power ratio (milliWatts for the phone example). This leads to confusion. SNR is frequently expressed as the log (base 10) of the ratio. When expressed as a log, the dimensionless unit of SNR is decibel (dB).

What is signal and what is noise can depend on the circumstances. Radio waves from lighting are noise to an AM radio broadcast, but can be signal to a meteorological experiment.

► Iris Device

Signature Benchmark

► Signature Databases and Evaluation

Signature Characteristics

► Signature Features

Signature Corporate

► Signature Databases and Evaluation

Signature Databases and Evaluation

Marcos Martinez-Diaz, Julian Fierrez
Biometric Recognition Group - ATVS,
Escuela Politecnica Superior, Universidad Autonoma de Madrid, Campus de Cantoblanco, Madrid, Spain

Synonyms

Signature benchmark; Signature corpora; Signature data set

Definition

Signature databases are structured sets of collected signatures from a group of individuals that are used either for evaluation of recognition algorithms or as part of an operational system.

Signature databases for evaluation purposes are, in general, collections of signatures acquired using a digitizing device such as a pen tablet or a touch-screen. Publicly available databases allow a fair performance comparison of signature recognition algorithms proposed by independent entities. Moreover, signature databases play a central role in public performance evaluations, which compare different recognition algorithms by using a common experimental framework. This type of databases is covered in this entry.

On the other hand, signature databases can also be a module of a verification or identification system.

They store signature data and other personal information of the enrolled users. This signature database is used during the operation of the recognition system to retrieve the enrolled data needed to perform the biometric matching. This kind of databases is not addressed here.

Dynamic Signature Databases

Until the beginning of this century, research on automatic signature verification had been carried out using privately collected databases, since no public ones were available. This fact limits the possibilities to compare the performance of different systems presented in the literature, which may have been tuned to specific capture conditions. Additionally, the usage of small data sets reduces the statistical relevance of experiments. The lack of publicly available databases has also been motivated by privacy and legal issues, although the data in these databases are detached from any personal information. The impact of the signature structural differences among cultures must also be taken into account when considering experimental results on a specific database. As an example, in Europe, signatures are usually formed by a fast writing followed by a flourish, while in North America, they usually correspond to the signers name with no flourish. On the other hand, signatures in Asia are commonly formed by Asian characters, which are composed of a larger number of short strokes compared with European or North American signatures.

While some authors have made public the databases used for their experimental results [1], most current dynamic signature databases are collected by the joint effort of different research institutions. These databases are, in general, freely available or can be obtained at a reduced cost. Many signature databases are part of larger multimodal biometric databases, which include other traits such as fingerprint or face data. This is done for two main reasons: the research interest on multimodal algorithms and the low effort required to incorporate the collection of other biometric traits once a database acquisition campaign has been organized.

Two main modalities in signature recognition exist. Off-line systems use signature images that have been previously captured with a scanner or camera. On the other hand, on-line systems employ digitized signals from the signature dynamics such as the pen position or pressure. These signals must be captured with specific devices such as ▶ pen tablets or ▶ touch-screens. The most popular databases in the biometric research community are oriented to on-line verification, although in some of them, the scanned signature images are also available [2, 3]. Some efforts have been carried out in the handwriting community to collect large off-line signature databases such as the GPDS-960 Corpus [4].

Unlike other biometric traits, signatures can be forged with relative ease. Signature verification systems must not only discriminate traits from different subjects (such as fingerprints) but also must discriminate between genuine signatures and forgeries. In general, signature databases provide a number of forgeries for the signatures of each user. The accuracy of the forgeries depends on the acquisition protocol, the skill of the forgers, and on how much time the forgers are let to train. Nevertheless, forgeries in signature databases are usually performed by subjects with no prior experience in forging signatures, this being a limitation to the quality of forgeries.

Most on-line signature databases have been captured with ▶ digitizing tablets. These tablets are, in general, based on an electromagnetic principle, allowing the detection of the pen position (x, y), inclination angles $(\theta, \gamma) = (\text{azimuth, altitude})$, and pressure p. They allow recording the pen dynamics even when the pen is not in contact with the signing surface (i.e., during pen-ups). On the other hand, databases captured with other devices such as touch-screens (e.g., PDAs) provide only pen position information, which is recorded exclusively when the pen is in contact with the device surface.

In the following, a brief description of the most relevant available on-line signature databases is given in chronological order.

PHILIPS Database

Signatures from 51 users were captured using a Philips Advanced Interactive Display (PAID) digitizing tablet at a sampling rate of 200 Hz [5]. The following signals were captured: position coordinates, pressure, azimuth, and altitude.

Each user contributed 30 genuine signatures, leading to 1,530 genuine signatures. Three types of forgeries are present in the database: 1,470 over-the-shoulder forgeries, 1,530 home-improved, and 240 professional

forgeries. There is not a fixed number of forgeries available for each user. Over-the-shoulder forgeries were produced by letting the forger observe the signing process. Home-improved forgeries were produced by giving to the forgers samples of the signature static image and letting them to practice at home. Professional forgeries were performed by forensic document examiners.

MCYT Bimodal Database

The MCYT bimodal database is comprised of signatures and fingerprints from 330 individuals [2]. Signatures were acquired using a Wacom Intuos A6 tablet with a sampling frequency of 100 Hz. The users signed repeatedly on a paper with a printed grid placed over the pen tablet. The following time sequences are captured: position coordinates, pressure, azimuth, and altitude.

There are 25 genuine signatures and 25 forgeries per user, leading to 16,500 signatures in the database. For each user, signatures were captured in groups of 5. First, 5 genuine signatures, then 5 forgeries from another user, repeating this sequence until 25 signatures from each type, were performed. Each user provided 5 forgeries for the 5 previous users in the database. As the user is forced to concentrate on different tasks between each group of genuine signatures, the variability between groups is expected to be higher than the one within the same group.

Genuine signatures and forgeries corresponding to 75 users from the MCYT database were scanned and are also available as an off-line signature database. This signature corpus is one of the most popular for the evaluation of signature verification algorithms that are being used by more than 50 research groups worldwide.

BIOMET Multimodal Database

The BIOMET multimodal database [6] is comprised of five modalities: audio (voice), face, hand, fingerprint, and signature. The signatures were captured using a Wacom Intuos2 A6 pen tablet and an ink pen with a sampling rate of 100 Hz. The pen coordinates, pen-pressure, azimuth, and altitude signals were captured. The database contains data from 84 users, with 15 genuine signatures and up to 12 forgeries per user. Signatures were captured in two sessions separated by 3–5 months. In the first session, 5 genuine signatures and 6 forgeries were acquired. The remaining 10 genuine signatures and 6 forgeries were captured in the second session. Forgeries are performed by 4 different users (3 forgeries each). This database contains 2,201 signatures, since not all users have complete data: 8 genuine signatures and 54 forgeries are missing.

SVC2004 Database

Two signature databases were released prior to the Signature Verification Competition (SVC) 2004 [7] for algorithm development and tuning. They were captured using a Wacom Intuos digitizing tablet and a Grip Pen. Due to privacy issues, users were advised to use invented signatures as genuine ones. Nevertheless, users were asked to thoroughly practice their invented signatures to reach a reasonable level of spatial and temporal consistency.

The two databases differ in the available data, and correspond to the two tasks defined in the competition. One contains only pen position information, while the other provides pressure and pen orientation (azimuth and altitude) signals also. Each database contains 40 users, with 20 genuine signatures and 20 forgeries per user acquired in two sessions, leading to 1,600 signatures per database. Forgeries for each user were produced by at least four other users, aided by a visual tool, which represented the signature dynamics on a display. Both occidental and asian signatures are present in the databases.

SUSIG Database

The SUSIG database consists of two sets: one captured using a pen tablet without visual feedback (Blind subcorpus) and the other using a pen tablet with an LCD display (Visual subcorpus) [8]. There are 100 users per database, but these do not coincide, as the Visual subcorpus was captured 4 years after the Blind one. For the Blind subcorpus, a WACOM Graphire2 pen tablet was used. The Visual subcorpus was acquired using an Interlink Electronics ePad-ink tablet, with a pressure-sensitive LCD. In both subcorpora, the pen coordinates and the pen pressure signals were captured using a sampling frequency of 100 Hz. While performing forgeries, users had prior visual input of the signing process on a separate screen or on the LCD display for the Blind and Visual subcorpus respectively.

For the Blind subcorpus, 8 or 10 genuine signatures were captured in a single session. The users also provided 10 forgeries from another randomly selected user. Two sessions were performed in the Visual subcorpus. During each one, users provided 10 genuine signatures and 5 forgeries.

MyIDea Multimodal Database

This signature set is a subset of the MyIDea Multimodal Biometric Database [9]. A Wacom Intuos2 A4 graphic tablet was used at a sampling rate of 100 Hz. Pen position, pressure, azimuth, and altitude signals were captured. This data set has the particularity that the user must read loud what he is writing, allowing what the authors call CHASM (Combined Handwriting and Speech Modalities). This corpus consists of ca. 70 users. Signatures were captured in 3 sessions. During each session, each user performed 6 genuine signatures and 6 forgeries, with visual access to the images of the target signatures.

BiosecurID Multimodal Database

This database was collected by 6 different Spanish research institutions [3]. It includes the following biometric traits: speech, iris, face, signature, handwriting, fingerprints, hand, and keystroke. The data were captured in 4 sessions distributed in a 4 month time span. The user population was specifically selected to contain a uniform distribution of users from different ages and genders. Nonbiometric data were also stored, such as age, gender, handedness, vision aids, and manual worker (if the user has eroded fingerprints). This allows studying specific demographic groups.

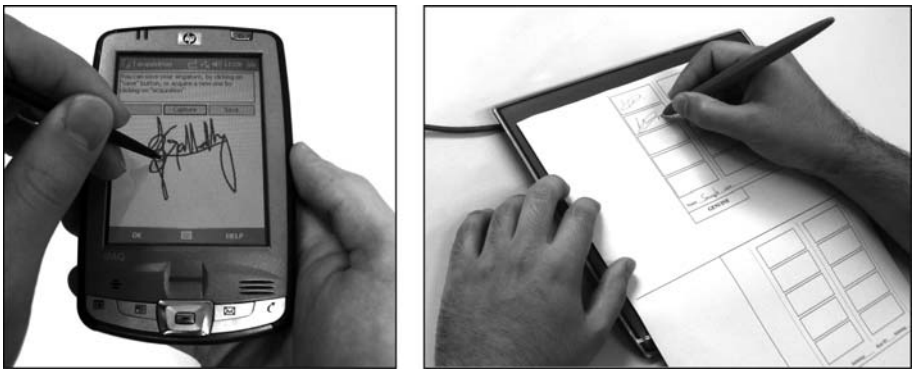
The signature pen-position, pressure, azimuth, and altitude signals were acquired using a Wacom Intuos3 A4 digitizer at 100 Hz. During each session, two signatures were captured at the beginning and two at the end, leading to 16 genuine signatures per user. Each user performed one forgery per session of signatures from other three users in the database. The skill level of the forgeries is increased by showing to the forger more information of the target signature incrementally. In the first session, forgers have only visual access to one genuine signature; more data (i.e., signature dynamics) are shown in further sessions and forgers are let more time to train. Off-line signature data are also available, since signatures were captured using an inking pen.

BioSecure Multimodal Database

The BioSecure Multimodal Database was collected by 11 European institutions under the BioSecure Network of Excellence [10]. It has three data sets captured in different scenarios: DS1 was captured remotely over the internet, DS2 was acquired in a desktop environment, and DS3 under mobile conditions. The database covers face, fingerprint, hand, iris, signature, and speech modalities and includes two signature subcorpora corresponding to the DS2 and DS3 data sets. These two data sets were produced by the same group of 667 users. The DS2 data set was captured using a Wacom Intuos3 A6 digitizer at 100 Hz and an ink pen while the user was sitting. On the other hand, the DS3 data set was captured with a PDA. Users were asked to sign while standing and holding the PDA in one hand, emulating realistic operating conditions. An HP iPAQ hx2790 with a sampling frequency of 100 Hz was used as capture device. The pen position, pressure, azimuth, and altitude signals are available in DS2, while only the pen position is available on DS3 due to the nature of the PDA touch-screen.

Signatures were captured in two sessions and in blocks of 5. An average of two months was left between each session. During each session, users were asked to perform 3 sets of 5 genuine signatures and 5 forgeries between each set. Following this protocol, each user performed 5 forgeries for the previous 4 users in the database. Thus, 30 genuine signatures and 20 forgeries are available for each user. Forgeries are collected in a "worst case" scenario. For DS2, the users had visual access to the dynamics of the signing process of the signatures they had to forge on a computer screen. In DS3, each forger had access to the dynamics of the genuine signature on the PDA screen and a tracker tool allowing to see the original strokes. Some users were even allowed to sign following the strokes produced by the tracker tool, reproducing thus the geometry and dynamics of the forged signature with high accuracy.

The DS3 data set is the first multisession database captured on a PDA and represents a very challenging database [11]. Apart from the high accuracy of the



Signature Databases and Evaluation. Figure 1 PDA signature capture process in the BioSecure DS3 - Mobile Scenario dataset (left) and pen-tablet capture process in the BioSecure DS2 - Access Control Scenario dataset (right). The acquisition setup and paper template used in DS2 is similar to the ones used in MCYT, BIOMET, MyIDea and BiosecurID.

forgeries, signatures from DS3 present sampling errors and irregular sampling rates. Moreover, pen position signals during pen-ups are not available, since the acquisition device captures the pen dynamics only when the PDA stylus is in contact with the touchscreen surface.

The capture process for both DS2 and DS3 is shown in Fig. 1. Examples of signatures from the BioSecure Signature subcorpora corresponding to DS2 and DS3 are presented in Fig. 2. Unconnected samples represent that at least one sample is missing between them due to sampling errors.

In Table 1, the main features of the described signature databases are presented.

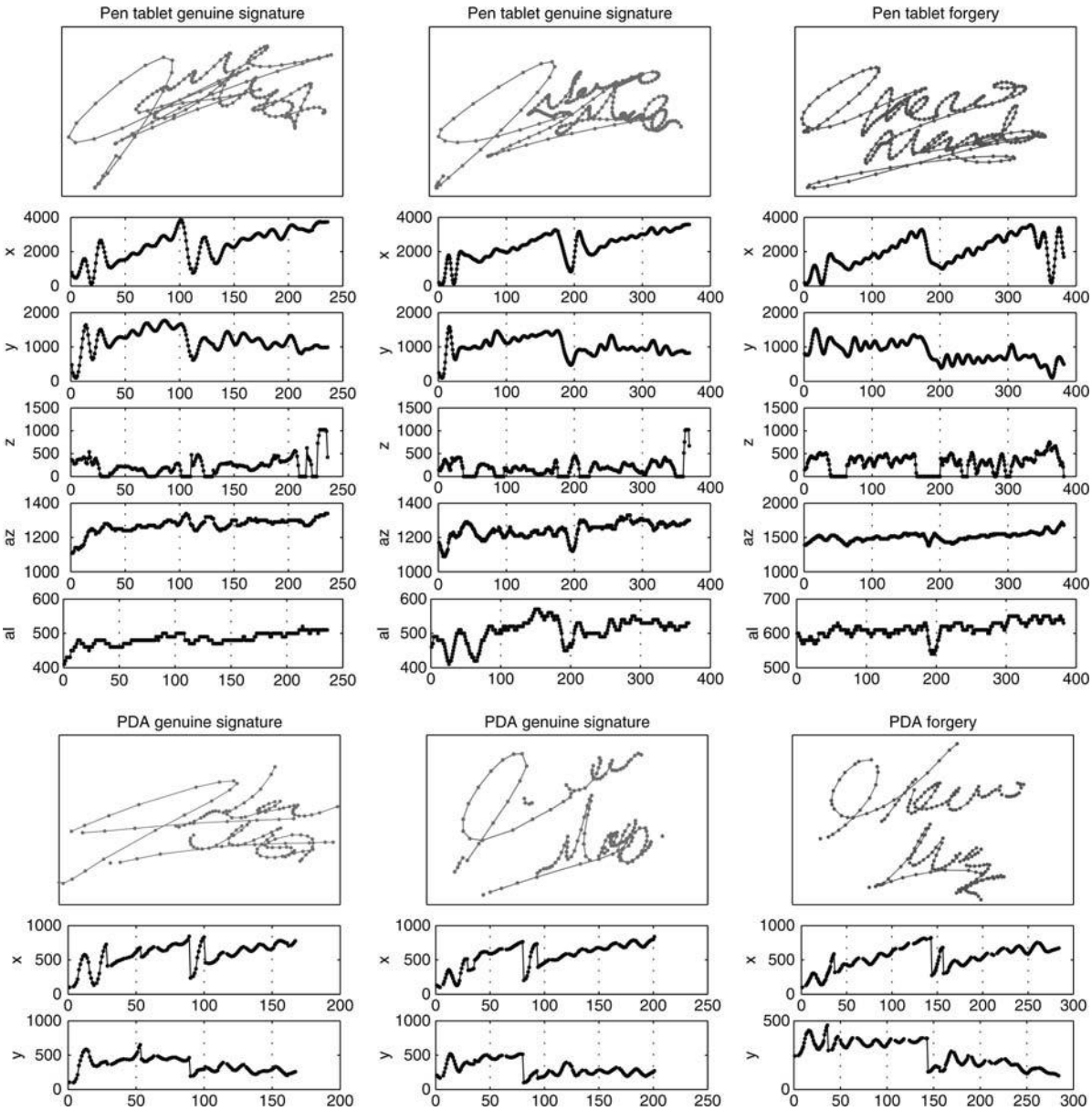
Signature Verification Evaluation Campaigns

Despite the usage of a common database, one of the main difficulties when comparing the performance of different biometric systems is the different experimental conditions, under which each system is evaluated by its designers. To overcome these difficulties, evaluations and competitions provide a common reference for system comparison on the same database and protocol. Public evaluations in the field of automatic signature verification are less common than for other biometric modalities such as fingerprint or speech. In particular, only evaluations for the on-line signature verification modality have been proposed. These include the Signature Verification Competition (SVC), which took place in 2004 [7], the signature modality of the BioSecure Multomodal Evaluation Campaign held in 2007 [12], and the BioSecure Signature Evaluation Campaign in 2009 [13].

Signature Verification Competition (SVC 2004)

The Signature Verification Competition (SVC 2004) represents the first public evaluation campaign in the field of signature verification [7]. The competition was divided into two tasks, depending on the available signature signals. In Task 1, only the pen position signals (x,y) and the sample timestamps were available. In Task 2, the pen pressure p and azimuth and altitude angles (θ,γ) were also available. Participants had prior access to a signature dataset for each task. These data sets were later released for public access, and are referred to as the SVC2004 database. Signatures from 40 users are present in each data set. This evaluation has the particularity that users were advised to use invented signatures because of privacy issues. Moreover, they did not have visual feedback from the signing process, since signatures were captured with a digitizing tablet and a special pen.

The evaluation results were first released to participants, which then had the choice to remain anonymous. The best Equal Error Rate (EER) in Task 1 was of 2.84% against ▶ skilled forgeries and 1.85% for ▶ random forgeries. In Task 2 (which included pressure and pen-inclination signals), the lowest EERs were 2.89% against skilled forgeries and 1.70% against random forgeries.



Signature Databases and Evaluation. Figure 2 Examples of signatures and associated signals from the BioSecure Multimodal Database DS2 and DS3 signature subcorpora captured using a pen tablet (top) and a PDA (bottom), respectively. As can be seen, there are missing samples for the signature captured with PDA, and no signals are available during pen-ups, contrary to the pen-tablet case.

BioSecure Multimodal Evaluation Campaign (BMEC 2007)

The BioSecure Multimodal Evaluation Campaign (BMEC) was held in 2007 with the aim of comparing the performance of verification systems from different research groups on individual biometric modalities and fusion scenarios [14]. Two scenarios were considered: access control and mobile conditions. In particular, the Mobile Scenario consisted of four modalities and fusion, using a subset of the BioSecure Multimodal Database DS3 captured on mobile conditions (i.e., using portable devices such as a PDA).

Signature Databases and Evaluation. **Table 1** Summary of the most popular on-line signature databases. The symbols x,y,p,θ,γ denote pen position horizontal coordinate, vertical coordinate, pen pressure, azimuth and altitude respectively

Name	Device	Users	Sessions	Signatures per user		Signals	Interval between sessions
				Genuine	Forgeries		
PHILIPS	Pen tablet	51	3–5	30	up to 70	x,y,p,θ,γ	1 week approx.
BIOMET	Pen tablet	84	3	15	up to 12	x,y,p,θ,γ	3–5 months
MCYT	Pen tablet	330	1	25	25	x,y,p,θ,γ	-
SVC2004 Task 1	Pen tablet	40	2	20	20	x,y	min. 1 week
SVC2004 Task 2	Pen tablet	40	2	20	20	x,y,p,θ,γ	min. 1 week
SUSIG Blind Subcorpus	Pen tablet	100	1	8 or 10	10	x,y,p	-
SUSIG Visual Subcorpus	Pen tablet	100	2	20	10	x,y,p	1 week approx.
MyIDea	Pen tablet	ca. 100	3	18	18	x,y,p,θ,γ	days to months
BioSecurID	Pen tablet	400	4	16	16	x,y,p,θ,γ	1 month approx.
BioSecure DS2	Pen tablet	ca. 650	2	30	20	x,y,p,θ,γ	1 month approx.
BioSecure DS3	PDA	ca. 650	2	30	20	x,y,p,θ,γ	1 month approx.

In this evaluation, a signature subset from the BioSecure Multimodal DS3 database was used. A set of signatures from 50 users was previously released to participants for algorithm development and tuning. For each user, 20 genuine signatures (15 from the first session and 5 from the second) as well as 20 forgeries were available.

Eleven signature verification systems from seven independent European research institutions were presented to the evaluation. The results of the evaluation and a description of each system that participated can be found in [12]. Another evaluation study in similar conditions, including a comparative analysis with respect to the BMEC participants, can be found in [11]. The best Equal Error Rate (EER) in the evaluation was of 4.03% for random forgeries and of 13.43% for skilled forgeries. The relatively high EER for skilled forgeries reveals the high quality of the forgeries acquired in this database.

BioSecure Signature Evaluation Campaign (BSEC 2009)

The BioSecure Signature Evaluation Campaign is aimed at measuring the impact of mobile acquisition conditions, time variability, and the information content of signatures in the performance of verification algorithms [13]. Signature subsets from the BioSecure Multimodal Databases DS2 (pen tablet) and DS3 (PDA touch-screen) corresponding to 50 users have been released to participants prior to the evaluation. At the time of publication, the results of the evaluation campaign are still not available.

Related Entries

- ▶ Biometric Sample Acquisition
- ▶ Off-line signature verification
- ▶ Performance Evaluation, Overview
- ▶ Signature Recognition

References

1. Munich, M.E., Perona, P.: Visual identification by signature tracking. IEEE Trans. Pattern Anal. Mach. Intell. **25**(2), 200–217 (2003)
2. Ortega-Garcia, J., Fierrez-Aguilar, et al.: MCYT baseline corpus: a bimodal biometric database. IEE Proc. Vis. Image Signal Process. **150**(6), 391–401 (2003)
3. Fierrez, J., Galbally, J., Ortega-Garcia, J., Freire, M.R., Alonso-Fernandez, F., Ramos, D., Toledano, D.T., Gonzalez-Rodriguez, J., Siguenza, J.A., Garrido-Salas, J., Anguiano-Rey, E., de Rivera, G.G., Ribalda, R., Faundez-Zanuy, M., Ortega, J.A., Cardenoso-Payo, V., Viloria, A., Vivaracho, C.E., Moro, Q.I., Igarza, J.J., Sanchez, J., Hernaez, I., Orrite-Urunuela, C., Martinez-Contreras, F., Gracia-Roche, J.J.: Biosecurid: A multimodal biometric database. Pattern Analysis & Applications (to appear) (2009)

4. Vargas, J., Ferrer, M., Travieso, C., Alonso, J.: Off-line handwritten signature GPDS-960 corpus. In: Proceedings of ninth International Conference on Document Analysis and Recognition, ICDAR, vol. 2, pp. 764–768. Curituba, Brazil (2007)
5. Dolfing, J.G.A., Aarts, E.H.L., van Oosterhout, J.J.G.M.: On-line signature verification with Hidden Markov Models. In: Proceedings of the International Conference on Pattern Recognition, ICPR, pp. 1309–1312. IEEE CS Press. Brisbane, Australia (1998)
6. Garcia-Salicetti, S., Beumier, C., Chollet, G., Dorizzi, B., Jardins, J.L.L., Lanter, J., Ni, Y., Petrovska-Delacretaz, D.: BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. In: Proceedings of IAPR International Conference on Audio- and Video-based Person Authentication, AVBPA, pp. 845–853. Springer LNCS-2688. Brisbane, Australia (2003)
7. Yeung, D.Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G.: SVC2004: First international signature verification competition. In: Proceedings of International Conference on Biometric Authentication, ICBA, pp. 16–22. Springer LNCS-3072 (2004)
8. Kholmatov, A., Yanikoglu, B.: Newblock Susig: an on-line signature database, associated protocols and benchmark results. Pattern Analysis & Applications (2008)
9. Dumas, B., Pugin, C., Hennebert, J., Petrovska-Delacretaz, D., Humm, A., Evequoz, F., Ingold, R., Rotz, D.V.: MyIDea - multimodal biometrics database, description of acquisition protocols. In: Proceedings of third COST 275 Workshop (COST 275), pp. 59–62. Hatfield, UK (2005)
10. Association BioSecure: BioSecure multimodal database. (http://www.biosecure.info) (2007). Last Accessed 03 March, 2009
11. Martinez-Diaz, M., Fierrez, J., Galbally, J., Ortega-Garcia, J.: Towards mobile authentication using dynamic signature verification: useful features and performance evaluation. In: Proc. Intl. Conf. on Pattern Recognition, ICPR pp. 1–6 (2008)
12. TELECOM & Management SudParis: BioSecure Multimodal Evaluation Campaign 2007 Mobile Scenario - experimental results. Tech. rep. (2007). (http://biometrics.it-sudparis.eu/BMEC2007/files/Results_mobile.pdf). Last Accessed 03 March, 2009
13. TELECOM & Management SudParis: Biosecure Signature Evaluation Campaign, BSEC 2009. http://biometrics.it-sudparis.eu/BSEC2009. URL http://biometrics.it-sudparis.eu/BSEC2009
14. Alonso-Fernandez, F., Fierrez, J., Ramos, D., Ortega-Garcia, J.: Dealing with sensor interoperability in multi-biometrics: the UPM experience at the BioSecure Multimodal Evaluation 2007. In: Defense and Security Symposium, Biometric Technologies for Human Identification, BTHI, Proc. SPIE, vol. 6944. Orlando, USA (2008)

Signature Dataset

▶ Signature Databases and Evaluation

Signature Features

Marcos Martinez-Diaz[1], Julian Fierrez[1], Seiichiro Hangai[2]
[1]Biometric Recognition Group - ATVS, Escuela Politecnica Superior, Universidad Autonoma de Madrid, Campus de Cantoblanco, Madrid, Spain
[2]Department of Electrical Engineering, Tokyo University of Science, Japan

Synonyms

Signature characteristics

Definition

Signature features represent magnitudes that are extracted from digitized signature samples, with the aim of describing each signature as a vector of values. The extraction and selection of optimum signature features is a crucial step when designing a verification system. Features must allow each signature to be described in a way that the discriminative power between signatures produced by different users is maximized while allowing variability among signatures from the same user.

On-line signature features can be divided into two main types. Global features model the signature as a holistic multidimensional vector and represent magnitudes such as average speed, total duration, and aspect ratio. On the other hand, local features are time-functions derived from the signals, such as the pen-position coordinate sequence or pressure signals, captured with digitizer tablets or touch-screens.

In off-line signature verification systems, features are extracted from a static signature image. They can also be classified as global, if they consider the image as a whole (e.g., image histogram, signature aspect ratio); or local, if they are obtained from smaller image regions (e.g., local orientation histograms).

This entry is focused on on-line signature features, although a brief outline of off-line signature features is also given.

Introduction

Several approaches to extract discriminative information from on-line signature data have been proposed

in the literature [1]. The existing systems can be broadly divided into two main types: *Global systems*, in which a holistic vector representation consisting of a set of global features (e.g., signature duration, direction after first pen-up) is derived from the signature trajectories [2, 3], and *function-based* systems, in which time sequences describing *local* properties of the signature are used for recognition [4, 5], (e.g., position, acceleration). Although recent works show that global approaches are competitive with respect to local methods in some circumstances [6], the latter approach has traditionally yielded better results. Despite this advantage, systems based on local features usually employ matching algorithms, which are computationally more expensive than global-feature ones.

Due to the usually low amount of training data in signature verification, ► feature selection techniques must be applied in order to reduce the feature vector dimensionality. These techniques allow of finding the optimal feature set for each system or scenario [7].

Feature extraction and preprocessing

Signature features are, in general, extracted from the time functions captured from the pen dynamics while an individual signs. In most cases, the capture of time functions from the handwritten signature is carried out with acquisition devices such as digitizing tablets or touch-screens. These devices provide pen position information (i.e. horizontal x and vertical y coordinates), and in some cases, pen pressure z and pen inclination (► azimuth and ► altitude). A diagram showing the nature of the captured signals and an example of the signals from a real signature are shown in Fig. 1. Other less common examples of on-line signature acquisition devices are special pens with dedicated hardware inside that captures signature data such as coordinate, force, or velocity information.

The sampling rate of these devices is, in general, between 100 and 200 Hz. Since the maximum frequencies of the pen movements during handwriting are 20-30 Hz [1], these sampling rates satisfy the Nyquist criterion.

Preprocessing steps before feature extraction may be performed, such as position, size and rotation normalization, noise filtering, or resampling. In some works, resampling is avoided as it degrades the velocity related features [4].

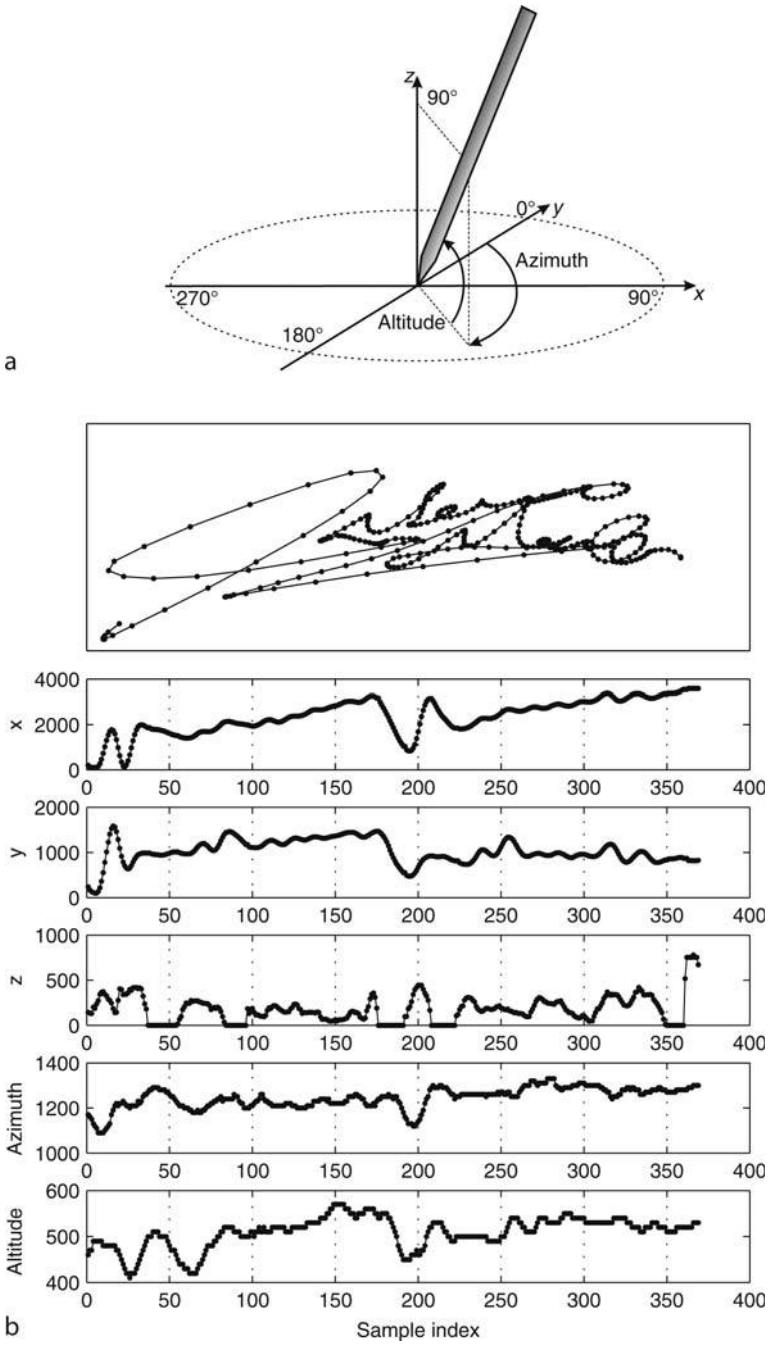
Global features

Global feature-based systems describe each signature as a multidimensional vector where each element consists on a feature extracted from the whole pen trajectory. Many feature sets have been proposed in the literature [2, 3, 8, 9], with variable sizes and a maximum size of 100 features [6]. Due to the training data scarcity and adverse effects of the curse of dimensionality, feature selection techniques must be applied to reduce the feature vector size. In Table 1, the 100 features described in [6] are presented. This global feature set includes most of the features described in previous works from other authors. Features are arranged in the order of descending individual discriminative power. In Fig. 2, examples of the distribution of global features presented in Table 1 are shown.

Local features

Local features represent time sequences extracted from the signature raw captured data. A set of local features leads to a multidimensional discrete sequence that describes a signature. Depending on the matching algorithm, feature sets of varying sizes have been proposed in the literature. As a rule of thumb, Dynamic Time Warping-based algorithms employ few local features, while systems based on Hidden Markov Models or Gaussian Mixture Models employ larger feature sets. In Table 2, the most popular local features found in the literature are presented [2, 3, 4, 5, 10, 11, 12].

As in the case of global features, feature selection algorithms must be applied to discriminate the best performing feature set. Usually, small feature sets are selected for Dynamic Time Warping-based matching algorithms. In these systems, speed-related features extracted from the first derivative of the pen-coordinate time sequences (features 10-11 in Table 2) have shown to be very effective [4]. On the other hand, larger feature sets are used when Hidden Markov or Gaussian Mixture Models are employed [5, 11] for signature matching. Features related to second-order derivatives (features 19-27 in Table 2) have not proved to significantly improve the system verification performance [3]. Examples of the local features presented in Table 2 are depicted in Fig. 3.



Signature Features. Figure 1 **(a)** Representation of the position, azimuth and altitude of the pen with respect to the capture device. **(b)** Example of raw captured data from a signature.

The usage of features related to pen orientation (azimuth and altitude) is a subject of controversy. Although some authors report that these features increase the verification performance [12], others have reported a low discriminative power for these features [2]. Moreover, these features are not always available, since many touch-screen acquisition devices such as Tablet-PCs or PDAs are unable to capture pen orientation information.

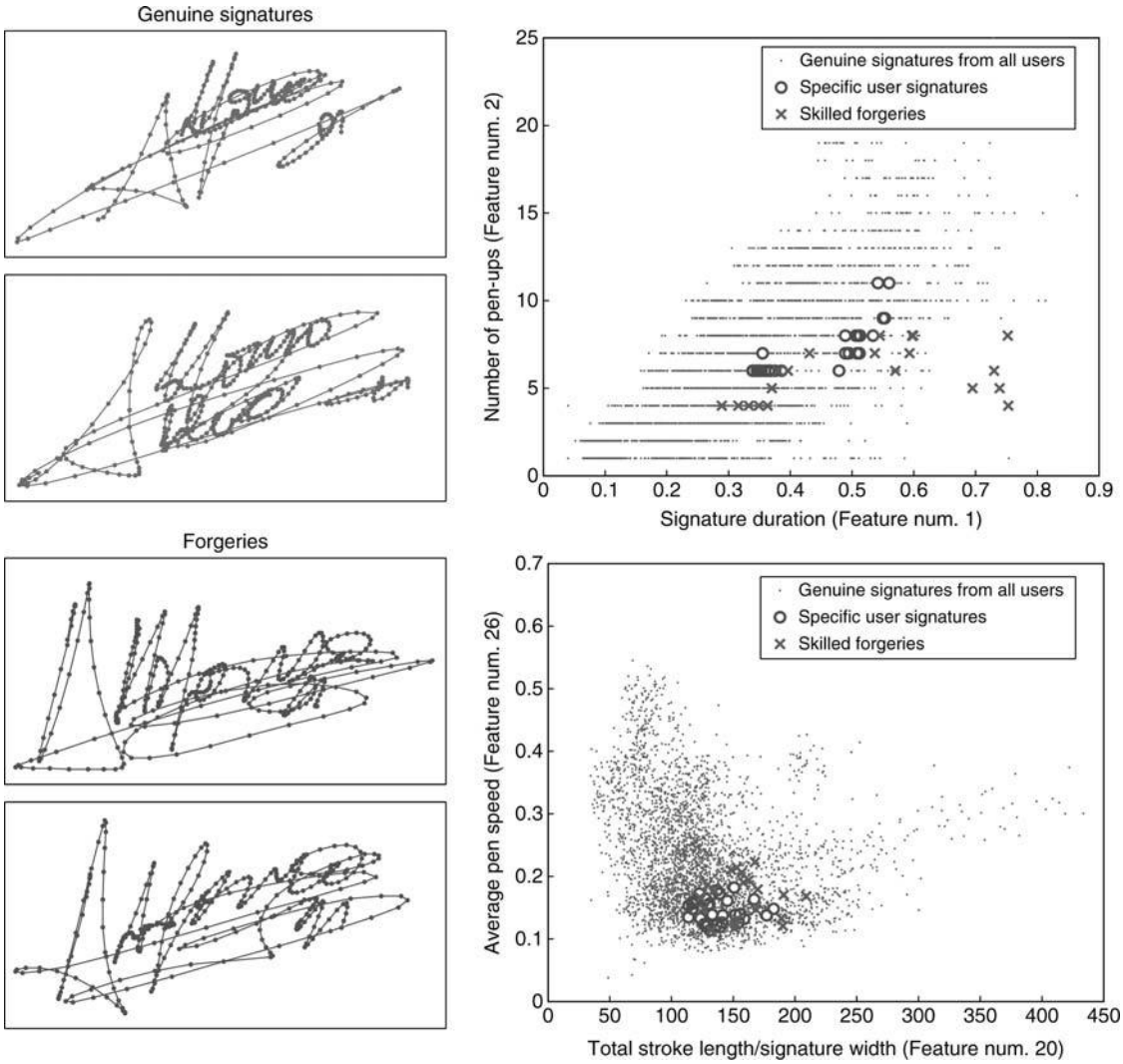
The fusion of the global and local feature-based systems has been reported to provide better performance than the individual systems [6].

Signature Features. **Table 1** Set of global features sorted by individual discriminative power (T denotes time interval, t denotes time instant, N denotes number of events, θ denotes angle. Note that some symbols are defined in different features of the table (e.g., Δ in feature 7 is defined in feature 15)

Ranking	Feature Description	Ranking	Feature Description
1	signature total duration T_s	2	N(pen-ups)
3	N(sign changes of dx/dt and dy/dt)	4	average jerk $\bar{j}$
5	standard deviation of a_y	6	standard deviation of v_y
7	(standard deviation of y)/Δ_y	8	N(local maxima in x)
9	standard deviation of a_x	10	standard deviation of v_x
11	j_{rms}	12	N(local maxima in y)
13	t(2nd pen-down)/T_s	14	(average velocity $\bar{v}$)/$v_{x,max}$
15	$\frac{A_{\min}=(y_{\max}-y_{\min})(x_{\max}-x_{\min})}{(\Delta_x=\sum_{i=1}^{\text{pen-downs}}(x_{\max\,i}-x_{\min\,i}))\Delta_y}$	16	$(x_{\text{last pen-up}}-x_{\max})/\Delta_x$
17	$(x_{\text{1st pen-down}}-x_{\min})/\Delta_x$	18	$(y_{\text{last pen-up}}-y_{\min})/\Delta_y$
19	$(y_{\text{1st pen-down}}-y_{\min})/\Delta_y$	20	$(T_w\bar{v})/(y_{\max}-y_{\min})$
21	$(T_w\bar{v})/(x_{\max}-x_{\min})$	22	(pen-down duration T_w)/T_s
23	$\bar{v}/v_{y,max}$	24	$(y_{\text{last pen-up}}-y_{\max})/\Delta_y$
25	$\frac{T((dy/dt)/(dx/dt)>0)}{T((dy/dt)/(dx/dt)<0)}$	26	$\bar{v}/v_{max}$
27	$(y_{\text{1st pen-down}}-y_{\max})/\Delta_y$	28	$(x_{\text{last pen-up}}-x_{\min})/\Delta_x$
29	(velocity rms v)/v_{max}	30	$\frac{(x_{\max}-x_{\min})\Delta_y}{(y_{\max}-y_{\min})\Delta_x}$
31	(velocity correlation $v_{x,y}$)/v_{max}^2	32	$T(v_y>0\,\vert\,\text{pen-up})/T_w$
33	$N(v_x=0)$	34	direction histogram s_1
35	$(y_{\text{2nd local max}}-y_{\text{1st pen-down}})/\Delta_y$	36	$(x_{\max}-x_{\min})/x_{\text{acquisition range}}$
37	$(x_{\text{1st pen-down}}-x_{\max})/\Delta_x$	38	$T(curvature>Threshold_{curv})/T_w$
39	(integrated abs. centr. acc. a_{lc})/a_{max}	40	$T(v_x>0)/T_w$
41	$T(v_x<0\,\vert\,\text{pen-up})/T_w$	42	$T(v_x>0\,\vert\,\text{pen-up})/T_w$
43	$(x_{\text{3rd local max}}-x_{\text{1st pen-down}})/\Delta_x$	44	$N(v_y=0)$
45	(acceleration rms a)/a_{max}	46	(standard deviation of x)/Δ_x
47	$\frac{T((dx/dt)(dy/dt)>0)}{T((dx/dt)(dy/dt)<0)}$	48	(tangential acceleration rms a_t)/a_{max}
49	$(x_{\text{2nd local max}}-x_{\text{1st pen-down}})/\Delta_x$	50	$T(v_y<0\,\vert\,\text{pen-up})/T_w$
51	direction histogram s_2	52	t(3rd pen-down)/T_s
53	(max distance between points)/$A_{\min}$	54	$(y_{\text{3rd local max}}-y_{\text{1st pen-down}})/\Delta_y$
55	$(\bar{x}-x_{\min})/\bar{x}$	56	direction histogram s_5
57	direction histogram s_3	58	$T(v_x<0)/T_w$
59	$T(v_y>0)/T_w$	60	$T(v_y<0)/T_w$
61	direction histogram s_8	62	(1st $t(v_{x,min})$)/T_w
63	direction histogram s_6	64	T(1st pen-up)/T_w
65	spatial histogram t_4	66	direction histogram s_4
67	$(y_{\max}-y_{\min})/y_{\text{acquisition range}}$	68	(1st $t(v_{x,max})$)/T_w
69	(centripetal acceleration rms a_c)/a_{max}	70	spatial histogram t_1
71	θ(1st to 2nd pen-down)	72	θ(1st pen-down to 2nd pen-up)
73	direction histogram s_7	74	$t(j_{x,max})/T_w$
75	spatial histogram t_2	76	$j_{x,max}$
77	θ(1st pen-down to last pen-up)	78	θ(1st pen-down to 1st pen-up)
79	(1st $t(x_{max})$)/T_w	80	$\bar{j}_x$
81	T(2nd pen-up)/T_w	82	(1st $t(v_{max})$)/T_w

Signature Features. Table 1 (Continued)

Ranking	Feature Description	Ranking	Feature Description
83	$j_{y,max}$	84	θ(2nd pen-down to 2nd pen-up)
85	j_{max}	86	spatial histogram t_3
87	(1st $t(v_{y,min})$)/T_w	88	(2nd $t(x_{max})$)/T_w
89	(3rd $t(x_{max})$)/T_w	90	(1st $t(v_{y,max})$)/T_w
91	$t(j_{max})/T_w$	92	$t(j_{y,max})/T_w$
93	direction change histogram c_2	94	(3rd $t(y_{max})$)/T_w
95	direction change histogram c_4	96	$\bar{j}_y$
97	direction change histogram c_3	98	θ(initial direction)
99	θ(before last pen-up)	100	(2nd $t(y_{max})$)/T_w



Signature Features. Figure 2 Examples of genuine signatures and forgeries (left) and scatter plots of 4 different global features from the 100-feature set presented in Table 1 (right). The signatures belong to the BioSecure database and the Figure has been adapted from [13].

Signature Features. Table 2 Extended set of local features. The upper dot notation (e.g., $\dot{x}_n$) indicates time derivative

#	Feature	Description
1	*x*-coordinate	x_n
2	*y*-coordinate	y_n
3	Pen-pressure	z_n
4	Path-tangent angle	$\theta_n = \arctan(\dot{y}_n / \dot{x}_n)$
5	Path velocity magnitude	$\upsilon_n = \sqrt{\dot{y}_n + \dot{x}_n}$
6	Log curvature radius	$\rho_n = log(1/\kappa_n) = log(\upsilon_n / \dot{\theta}_n)$, where κ_n is the curvature of the position trajectory
7	Total acceleration magnitude	$a_n = \sqrt{t_n^2 + c_n^2} = \sqrt{\dot{\upsilon}_n^2 + \upsilon_n^2 \theta_n^2}$, where t_n and c_n are respectively the tangential and centripetal acceleration components of the pen motion
8	Pen azimuth	γ_n
9	Pen altitude	ϕ_n
10–18	First-order derivative of features 1–9	$\dot{x}_n, \dot{y}_n, \dot{z}_n, \dot{\theta}_n, \dot{\upsilon}_n, \dot{\rho}_n, \dot{a}_n, \dot{\gamma}_n, \dot{\phi}_n$
19–27	Second-order derivative of features 1–9	$\ddot{x}_n, \ddot{y}_n, \ddot{z}_n, \ddot{\theta}_n, \ddot{\upsilon}_n, \ddot{\rho}_n, \ddot{a}_n, \ddot{\gamma}_n, \ddot{\phi}_n$
28	Ratio of the minimum over the maximum speed over a window of 5 samples	$\upsilon_n^r = min\ \{\upsilon_{n-4}, \ldots, \upsilon_n\} / \max\ \{\upsilon_{n-4}, \ldots, \upsilon_n\}$
29–30	Angle of consecutive samples and first order difference	$\alpha_n = \arctan(y_n - y_{n-1} / x_n - x_{n-1})\ \dot{\alpha}_n$
31	Sine	$s_n = \sin(\alpha_n)$
32	Cosine	$c_n = \cos(\alpha_n)$
33	Stroke length to width ratio over a window of 5 samples	$r_n^5 = \frac{\sum_{k=n-4}^{k=n} \sqrt{(x_k - x_{k-1})^2 + (y_k - y_{k-1})^2}}{\max\{x_{n-4}, \ldots, x_n\} - \min\{x_{n-4}, \ldots, x_n\}}$
34	Stroke length to width ratio over a window of 7 samples	$r_n^7 = \frac{\sum_{k=n-6}^{k=n} \sqrt{(x_k - x_{k-1})^2 + (y_k - y_{k-1})^2}}{\max\{x_{n-6}, \ldots, x_n\} - \min\{x_{n-6}, \ldots, x_n\}}$

Off-line signature features

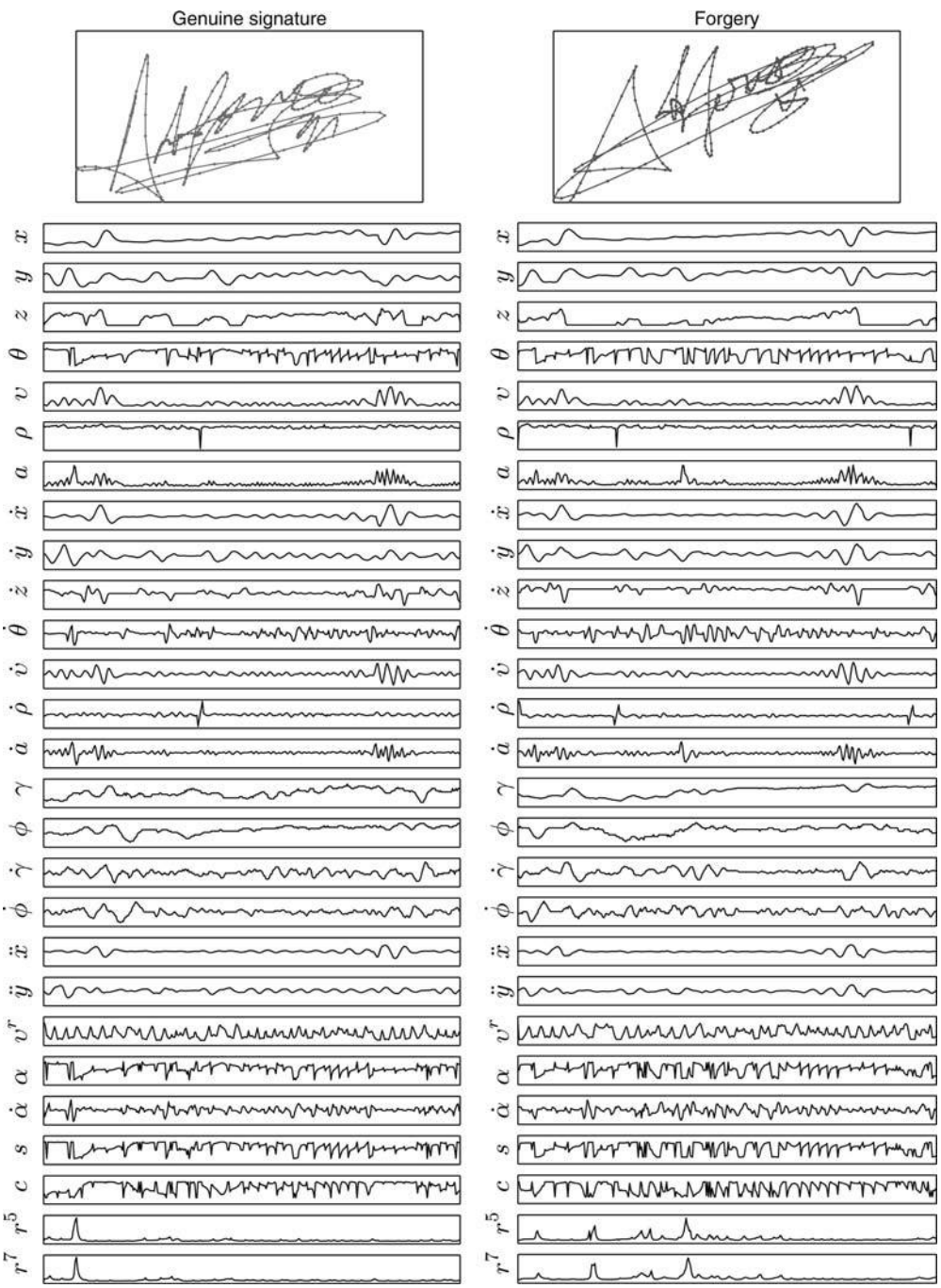
Off-line signature verification systems usually rely on image processing and shape recognition techniques to extract features. As a consequence, additional preprocessing steps such as image segmentation and binarization must be carried out. Features are extracted from gray-scale images, binarized images, or skeletonized images, among other possibilities. The proposed feature sets in the literature are notably heterogeneous, specially when compared with the case of on-line verification systems. These include, among others, the usage of image transforms (e.g., Hadamard), morphological operators, structural representations, ▶ graphometric features [14], directional histograms, and geometric features. Readers are referred to [15] for an exhaustive listing of off-line signature features.

Related Entries

- ▶ Feature Extraction
- ▶ Off-line Signature Verification
- ▶ On-line Signature Verification
- ▶ Signature Matching
- ▶ Signature Recognition

References

1. Plamondon, R., Lorette, G.: Automatic signature verification and writer identification: the state of the art. Pattern Recogn. **22**(2), 107–131 (1989)
2. Lei, H., Govindaraju, V.: A comparative study on the consistency of features in on-line signature verification. Pattern Recogn. Lett. **26**(15), 2483–2489 (2005)



Signature Features. Figure 3 Examples of functions from the 27-feature set presented in Table 2 for a genuine signature (left) and a forgery (right) of a particular subject.

3. Richiardi, J., Ketabdar, H., Drygajlo, A.: Local and global feature selection for on-line signature verification. In: Proceedings of IAPR eighth International Conference on Document Analysis and Recognition, ICDAR, Seoul, Korea (2005)
4. Kholmatov, A., Yanikoglu, B.: Identity authentication using improved online signature verification method. Pattern Recogn. Lett. **26**(15), 2400–2408 (2005)
5. Fierrez, J., Ramos-Castro, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: HMM-based on-line signature verification: feature extraction and signature modeling. Pattern Recogn. Lett. **28**(16), 2325–2334 (2007)
6. Fierrez-Aguilar, J., Nanni, L., Lopez-Penalba, J., Ortega-Garcia, J., Maltoni, D.: An on-line signature verification system based on fusion of local and global information. In: Proceedings of IAPR

International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Springer LNCS-3546, pp. 523–532 (2005)
7. Jain, A.K., Zongker, D.: Feature selection: evaluation, application, and small sample performance. IEEE Trans. Pattern Anal. Mach. Intell. **19**(2), 153–158 (1997)
8. Nelson, W., Turin, W., Hastie, T.: Statistical methods for on-line signature verification. Int. J. Pattern Recogn. Artif. Intell. **8**(3), 749–770 (1994)
9. Lee, L.L., Berger, T., Aviczer, E.: Reliable on-line human signature verification systems. IEEE Trans. Pattern Anal. Mach. Intell. **18**(6), 643–647 (1996)
10. Dolfing, J.G.A., Aarts, E.H.L., van Oosterhout, J.J.G.M.: On-line signature verification with Hidden Markov Models. In: Proceedings of the International Conference on Pattern Recognition, IEEE Press, pp. 1309–1312 (1998)
11. Van, B.L., Garcia-Salicetti, S., Dorizzi, B.: On using the Viterbi path along with HMM likelihood information for online signature verification. IEEE Trans. Syst. Man Cybern. B **37**(5), 1237–1247 (2007)
12. Muramatsu, D., Matsumoto, T.: Effectiveness of pen pressure, azimuth, and altitude features for online signature verification. In: Proceedings of IAPR International Conference on Biometrics, ICB, Springer LNCS 4642 (2007)
13. Martinez-Diaz, M., Fierrez, J., Galbally, J., Ortega-Garcia, J.: Towards mobile authentication using dynamic signature verification: useful features and performance evaluation. In: Proceedings International Conference on Pattern Recognition, ICPR, pp. 1–6 (2008)
14. Sabourin, R.: In: Off-line signature verification: recent advances and perspectives. Lect. Notes Comput. Sci. **1339** 84–98 (1997)
15. Impedovo, D., Pirlo, G.: Automatic signature verification: The state of the art. IEEE Trans. Syst. Man Cybern. C Appl. Rev. **38**(5), 609–635 (2008)

Signature Matching

Marcos Martinez-Diaz[1], Julian Fierrez[1], Seiichiro Hangai[2]
[1]Biometric Recognition Group - ATVS, Escuela Politecnica Superior, Universidad Autonoma de Madrid, Madrid, Spain
[2]Department of Electrical Engineering Tokyo University of Science, Japan

Synonyms

Signature similarity computation

Definition

The objective of signature matching techniques is to compute the similarity between a given signature and a signature model or reference signature set. Several pattern recognition techniques have been proposed as matching algorithms for signature recognition. In on-line signature verification systems, signature matching algorithms have followed two main approaches. Feature-based algorithms usually compute the similarity among multidimensional feature vectors extracted from the signature data with statistical classification techniques. On the other hand, function-based algorithms perform matching by computing the distance among time-sequences extracted from the signature data with technique such as Hidden Markov Models and Dynamic Time Warping. Off-line signature matching has followed many different approaches, most of which are related to image processing and shape recognition.

This essay focuses on on-line signature matching, although off-line signature matching algorithms are briefly outlined.

Introduction

As in other biometric modalities, signature matching techniques vary depending on the nature of the features that are extracted from the signature data. In *feature-based* systems (also known as global), each signature is represented as a multidimensional feature vector, while in function-based systems (also known as local) signatures are represented by multidimensional time sequences. Signature matching algorithms also depend on the enrollment phase. *Model-based* systems estimate a statistical model for each user from the training signature set. On the other hand, in *reference-based* systems the features extracted from the set of training signatures are stored as a set of template signatures. Consequently, given an input signature, in model-based systems the matching is performed against a statistical model, while in reference-based systems the input signature is compared with all the signatures available in the ► reference set.

Feature-Based Systems

Feature-based systems usually employ classical pattern classification techniques. In reference-based systems, the ► matching score is commonly obtained by using a distance measure between the feature vectors of input and template signatures [1, 2], or a trained classifier. Distance

measures used for signature matching include Eucli dean, weighted Euclidean, and Mahalanobis distance. In model-based systems, trained classifiers are employed, including approaches such as Neural Networks, Gaussian Mixture Models [3] or Parzen Windows [4].

Function-Based Systems

In these systems, multidimensional time sequences extracted from the signature dynamics are used as features. Given the similarity of this task to others related to speaker recognition, the most popular approaches in local signature verification are related to algorithms proposed in the speech recognition community.

Among these, signature verification systems using ▶ Dynamic Time Warping (DTW) [5, 6, 7] or Hidden Markov Models (HMM) [8, 9, 10, 11] are the most popular approaches in signature verification. In such systems, the captured time functions (e.g., pen coordinates, pressure, etc.) are used to model each user signature. In the following, Dynamic Time Warping and Hidden Markov Models are outlined. An brief overview of other techniques is also given.

Dynamic Time Warping

▶ Dynamic Time Warping (DTW) is an application of the Dynamic Programming principles to the problem of matching discrete time sequences. DTW was originally proposed for speech recognition applications [12]. The goal of DTW is to find an elastic match among samples of a pair of sequences X and Y that minimizes a predefined distance measure. The algorithm is described as follows. Let's define two sequences

$$\begin{aligned} X &= x_1, x_2, \ldots, x_i, \ldots, x_I \\ Y &= y_1, y_2, \ldots, y_j, \ldots, y_J \end{aligned} \tag{1}$$

and a distance measure as

$$d(i,j) = \left\| x_i - y_j \right\| \tag{2}$$

between sequence samples. A warping path can be defined as

$$C = c_1, c_2, \ldots, c_k, \ldots, c_K \tag{3}$$

where each c_k represents a correspondence (i, j) between samples of X and Y. The initial condition of the algorithm is set to

$$g_1 = g(1,1) = d(1,1) \cdot w(1) \tag{4}$$

where g_k represents the accumulated distance after k steps and $w(k)$ is a weighting factor that must be defined. For each iteration, g_k is computed as

$$g_k = g(i,j) = \min_{c_{k-1}} \left[g_{k-1} + d(c_k) \cdot w(k) \right] \tag{5}$$

until the Ith and Jth sample of both sequences respectively is reached. The resulting normalized distance is

$$D(X,Y) = \frac{g_K}{\sum_{k=1}^{K} w(k)} \tag{6}$$

where $\sum w(k)$ compensates the effect of the length of the sequences.

The weighting factors $w(k)$ are defined in order to restrict which correspondences among samples of both sequences are allowed. In Fig. 1a, a common definition of $w(k)$ is depicted, and an example of a warping path between two sequences is given. In this case, only three transitions are allowed in the computation of g_k. Consequently, Eq. (5) becomes

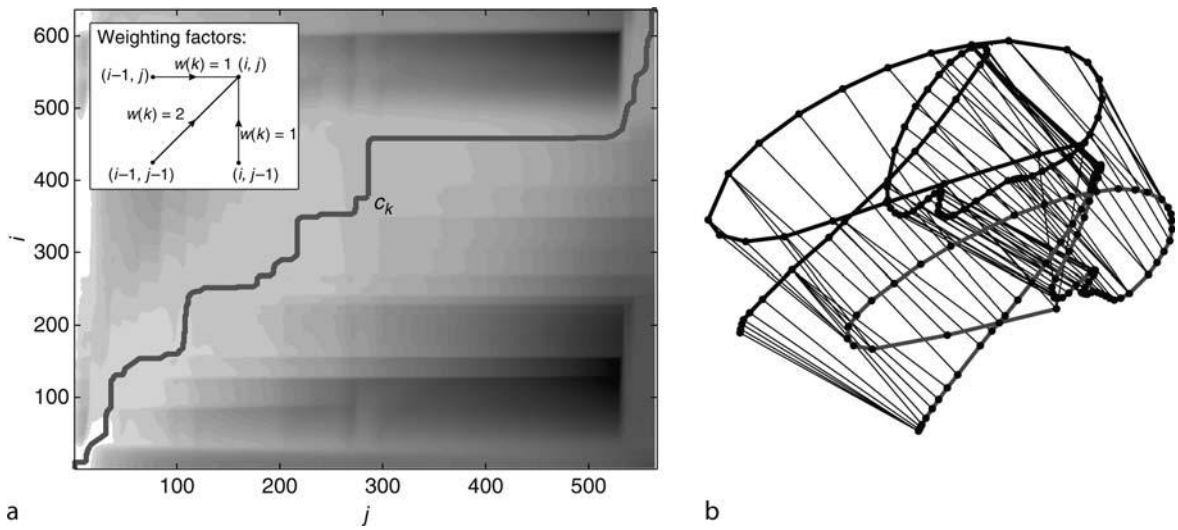
$$g_k = g(i,j) = \min \begin{bmatrix} g(i,j-1) + d(i,j) \\ g(i-1,j-1) + 2d(i,j) \\ g(i-1,j) + d(i,j) \end{bmatrix} \tag{7}$$

which is one of the most common implementations found in the literature. In Fig. 1b, an example of point correspondences between two signatures is depicted to visually show the results of the elastic alignment.

The algorithm has been further refined for signature verification by many authors [5, 7], reaching a notable verification performance. For example, the distance measure $d(i, j)$ can be alternatively defined, or other normalization techniques may be applied to the accumulated distance g_K among sequences. DTW can be also applied independently for each stroke, which may be specially well suited for oriental signatures, since they are generally composed of several strokes. Although the DTW algorithm has been replaced in speech-related applications by more powerful approaches such as HMMs, it remains as a highly effective tool for signature verification as it is best suited for small amounts of training data, which is the common case in signature verification.

Hidden Markov Models

Hidden Markov Models (HMM) have been widely used for speech recognition applications [13] as well



Signature Matching. Figure 1 (**a**) Optimal warping path between two sequences obtained with DTW. Point-to-point distances are represented with different shades of gray, lighter shades representing shorter distances and darker shades representing longer distances. (**b**) Example of point-to-point correspondences between two genuine signatures obtained using DTW.

as in many handwriting recognition applications. Several approaches using HMMs for dynamic signature verification have been proposed in the last years [8, 9, 10, 11]. An HMM represents a double stochastic process, governed by an underlying Markov chain, with a finite number of states and a random function set that generate symbols or observations each of which is associated with one state [11]. Observations in each state are modeled with GMMs in most speech and handwriting recognition applications. In fact, GMMs can be considered a single-state HMM and have also been successfully used for signature verification [14]. Given a sequence of multidimensional vectors of observations **O** defined as

$$\mathbf{O} = \mathbf{o}_1, \mathbf{o}_2, \ldots, \mathbf{o}_N,$$

corresponding to a given signature, the goal of HMM-based signature matching is to find the probability that this sequence has been produced by a Hidden Markov Model M

$$P(\mathbf{O}|M),$$

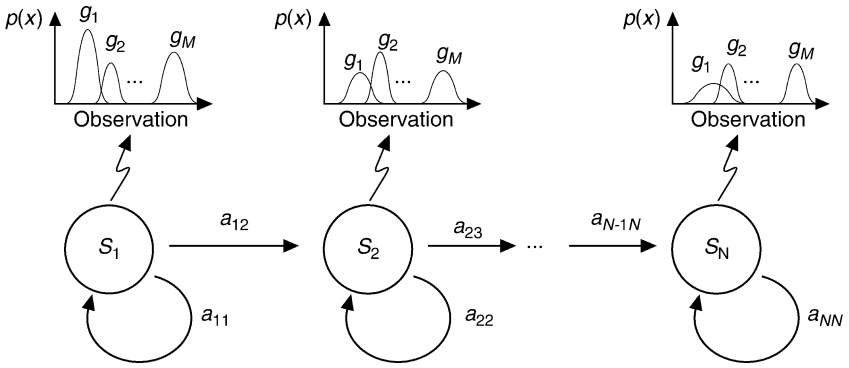
where M is the signature model computed during enrollment.

The basic structure of an HMM using GMMs to model observations is defined by the following elements:

- Number of hidden states N.
- Number of Gaussian mixtures per state M.
- Probability transition matrix $\mathbf{A} = \{a_{ij}\}$, which contains the probabilities of jumping from one state to another or staying on the same state.

In Fig. 2, an example of a possible HMM configuration is shown. Hidden Markov Models are usually trained in two steps using the enrollment signatures. First, state transition probabilities and observation statistical models are estimated using a Maximum Likelihood algorithm. After this, a re-estimation step is carried out using the Baum-Welch algorithm. The likelihood between a trained HMM and an input sequence (i.e., the matching score) is computed by using the Viterbi algorithm. In [10], the Viterbi path (that is, the most probable state transition sequence) is also used as a similarity measure. A detailed description of Hidden Markov Models is given in [13].

Within HMM-based dynamic signature verification, the existing approaches can be divided in *regional* and *local*. In regional approaches, the extracted time



Signature Matching. **Figure 2** Graphical representation of a left-to-right *N*-state HMM, with *M*-component GMMs representing observations and no skips between states.

sequences are further segmented and converted into a sequence of feature vectors or observations, each one representing regional properties of the signature signal [9, 11]. Some examples of segmentation boundaries are null vertical velocity points [9] or changes in the quantized trajectory direction [11]. On the other hand, local approaches directly use the time functions as observation sequences for the signature modeling [8, 10, 14].

Finding a reliable and robust model structure for dynamic signature verification is not a trivial task. While too simple HMMs may not allow to model properly the user signatures, too complex models may not be able to model future realizations due to overfitting. On the other hand, as simple models have less parameters to be estimated, their estimation may be more robust than for complex models. Two main parameters are commonly considered while selecting an optimal model structure: the number of states and the number of Gaussian mixtures per state [8]. Some approaches consider a user-specific number of states [10], proportional to the average signature duration or a user-specific number of mixtures [14]. Most of the proposed systems consider a left-ro-right configuration without skips between states, also known as Bakis topology (see Fig. 2).

Other Techniques

More examples of signature matching techniques include Neural Networks, in particular Bayesian, multilayer, time-delay Neural Networks and radial-basis functions among others have been applied for signature matching. Other examples include Structural approaches, which model signatures as a sequence, tree or graph of symbols. Support Vector Machines have also been applied for signature matching. The reader is referred to [15] for an exhaustive list of references related to these approaches.

Fusion of the feature- and function-based approaches has been reported to provide better performance than the individual systems [4].

Off-line Signature Matching

The proposed approaches for off-line signature matching are notably heterogeneous compared to on-line signature verification. These are mostly related to image and shape recognition techniques and classical statistical pattern recognition algorithms. They include Neural Networks, Hidden Markov Models, Support Vector Machines and distance-based classifiers among others. A summary of off-line signature matching techniques can be found in [15].

Related Entries

- ▶ Off-line Signature Verification
- ▶ On-line Signature Verification
- ▶ Signature Features
- ▶ Signature Recognition

References

1. Nelson, W., Turin, W., Hastie, T.: Statistical methods for on-line signature verification. Int. J. Pattern Recogn. Artif. Intell. **8**(3), 749–770 (1994)
2. Lee, L.L., Berger, T., Aviczer, E.: Reliable on-line human signature verification systems. IEEE Trans. Pattern Anal. Mach. Intell. **18**(6), 643–647 (1996)
3. Martinez-Diaz, M., Fierrez, J., Ortega-Garcia, J.: Universal Background Models for dynamic signature verification. In: Proceedings IEEE Conference on Biometrics: Theory, Applications and Systems, BTAS, pp. 1–6 (2007)
4. Fierrez-Aguilar, J., Nanni, L., Lopez-Penalba, J., Ortega-Garcia, J., Maltoni, D.: An on-line signature verification system based on fusion of local and global information. In: Proceedings of IAPR International Conference on Audio- and Video-Based Biometric Person Authentication, AVBPA, Springer LNCS-3546, pp. 523–532 (2005)
5. Sato, Y., Kogure, K.: Online signature verification based on shape, motion and writing pressure. In: Proceedings of sixth International Conference on Pattern Recognition, pp. 823–826 (1982)
6. Martens, R., Claesen, L.: Dynamic programming optimisation for on-line signature verification. In: Proceedings fourth International Conference on Document Analysis and Recognition, ICDAR, vol. 2, pp. 653–656 (1997)
7. Kholmatov, A., Yanikoglu, B.: Identity authentication using improved online signature verification method. Pattern Recogn. Lett. **26**(15), 2400–2408 (2005)
8. Fierrez, J., Ramos-Castro, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: HMM-based on-line signature verification: feature extraction and signature modeling. Pattern Recogn. Lett. **28**(16), 2325–2334 (2007)
9. Dolfing, J.G.A., Aarts, E.H.L., van Oosterhout, J.J.G.M.: On-line signature verification with Hidden Markov Models. In: Proceedings of the International Conference on Pattern Recognition, ICPR, pp. 1309–1312. IEEE CS Press (1998)
10. Van, B.L., Garcia-Salicetti, S., Dorizzi, B.: On using the Viterbi path along with HMM likelihood information for online signature verification. IEEE Trans. Syst. Man Cybern. B **37**(5), 1237–1247 (2007)
11. Yang, L., Widjaja, B.K., Prasad, R.: Application of Hidden Markov Models for signature verification. Pattern Recogn. **28**(2), 161–170 (1995)
12. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoust. **26**, 43–49 (1978)
13. Rabiner, L.R.: A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of the IEEE **77**(2), 257–286 (1989)
14. Richiardi, J., Drygajlo, A.: Gaussian Mixture Models for on-line signature verification. In: Proceedings of ACM SIGMM Workshop on Biometric Methods and Applications, WBMA. pp. 115–122 (2003)
15. Impedovo, D., Pirlo, G.: Automatic signature verification: The state of the art. IEEE Trans. Syst. Man. Cybern. C Appl. Rev. **38**(5), 609–635 (2008)

Signature Recognition

Olaf Henniger[1], Daigo Muramatsu[2], Takashi Matsumoto[3], Isao Yoshimura[4], Mitsu Yoshimura[5]

[1]Fraunhofer Institute for Secure Information Technology, Darmstadt, Germany

[2]Seikei University, Musashino-shi, Tokyo, Japan

[3]Waseda University, Shinjuku-ku, Tokyo, Japan

[4]Tokyo University of Science, Shinjuku-ku, Tokyo, Japan

[5]Ritsumeikan University, Sakyo-ku, Kyoto, Japan

Synonyms

Handwritten signature recognition; signature/sign recognition

Definition

A signature is a handwritten representation of name of a person. Writing a signature is the established method for authentication and for expressing deliberate decisions of the signer in many areas of life, such as banking or the conclusion of legal contracts. A related concept is a handwritten personal sign depicting something else than a person's name. As compared to text-independent writer recognition methods, signature/sign recognition goes with shorter handwriting probes, but requires to write the same name or personal sign every time. Handwritten signatures and personal signs belong to the behavioral biometric characteristics as the person must become active for signing.

Regarding the automated recognition by means of handwritten signatures, there is a distinction between on-line and off-line signature recognition. On-line signature data are captured using digitizing pen tablets, pen displays, touch screens, or special pens and include information about the pen movement over time (at least the coordinates of the pen tip and possibly also the pen-tip pressure or pen orientation angles over time). In this way, on-line signature data represent the way a signature is written, which is also referred to as signature dynamics. By contrast, off-line (or static) signatures are captured as grey-scale images using devices such as image scanners and lack temporal information.

Overview

Off-line and On-line Signatures

A number of features that are suitable for automated comparison can be extracted from handwritten signatures. The features depend on the type of data captured and the chosen comparison method. Handwriting data can be classified into

- on-line data captured during the writing process using devices such as digitizing tablets, Tablet PC's, or special pens
- off-line data captured from paper after the writing process using devices such as image scanners or cameras

Dynamic information, like ▶ stroke order, writing ▶ speed, and ▶ pen pressure, is available in on-line data, whereas only static information, like the shapes of handwritten characters, is available in off-line data. There is a multitude of methods for comparing handwritten signatures, see under on-line signatures and off-line signatures below.

Applications

Handwritten signatures are generally used for verification (confirming a claimed identity through one-to-one comparisons of biometric features), but rarely for identification (finding identifiers attributable to a person through one-to-many search among biometric features in a large database) [1]. Handwritten signatures have been used for a long time for authentication purposes in many applications, such as credit cards, banking transactions, agreements, and legal documents. Off-line signature serve as a unique means to verify the authenticity of a person through past records, such as signatures on traveler's cheques.

Strengths and Weaknesses

In order that handwritten signatures are useable for recognizing persons, genuine signatures (i.e. signatures written by the persons themselves whose names they represent) need to be sufficiently repeatable over time with respect to the comparison criteria, and forgeries (i.e. signatures not written by the persons themselves whose names they represent) need to be distinguishable from the genuine signatures by means of the comparison criteria [2]. As for some persons the handwritten signatures may vary considerably from signature to signature, the permanence of handwritten signatures is considered lower than that of many physiological biometric characteristics. As forgers can learn with some effort how to imitate the signatures of their victims, also the distinctiveness of handwritten signatures is lower than that of many physiological biometric characteristics. However, forging the signature dynamics is considerably harder than just forging the signature shape [3, 4] because information about the signature dynamics is less easily accessible to potential forgers than information about the signature shape.

The strengths of handwritten signatures compared to other biometric characteristics lie in a high level and wide spread of user acceptance and in the fact that handwritten signatures are regarded as an evidence of a deliberate decision on the part of the signer. Furthermore, people can modify their signatures in case of successful forgeries. By contrast, physiological biometric characteristics such as fingerprints or irises cannot be modified.

Performance Testing of Signature Verification Systems

Types of Testing

Users may obtain feedback (accepted or rejected) from the signature verification system, which then influences their subsequent input signatures. Genuine users will become accustomed to the system, and forgers will use this feedback to improve their signatures. Thus, in order to determine the performance in practice, scenario tests and operational tests (where users input signatures to a prototype or operational verification system) are preferable. Such tests, however, are expensive and time consuming, and it is extremely difficult to conduct tests on a large population. Moreover, since the decision threshold of the system must be fixed, we cannot evaluate the overall performance of the system.

Tests on databases can be easily conducted and a large population can be tested. Moreover, such tests can be conducted under the same conditions whenever needed, and we can compare the different systems and evaluate the overall performance of a system by assuming different decision thresholds. Thus, technology

tests using pre-existing or specially collected databases of genuine signatures and forgeries are useful to develop practical systems.

Neither scenario nor technology tests, by itself, are sufficient to evaluate the actual performance. Therefore, both tests should be conducted to estimate the performance of the system in practice.

Evaluation Protocol

The performance of a signature recognition system depends on the writing conditions and instruments, skill of writers, required level of similarity for acceptance or rejection, and precision and reliability of templates. The performance of a system is also highly dependent on the quality of the sample images to be verified. Thus, the design and configuration of the test should be carefully considered when performing evaluation experiments:

- The number of training signatures input during enrollment: 3–20 training signatures, or more, have generally been used in previous studies. For instance, in [5], five signatures were used for training.
- Selection method of training signatures from a database: Some researchers select training signatures randomly from genuine signatures, and some researchers select them from the first several signatures of the first session. Considering the actual situation, it is reasonable to select the first several signatures for training. In [5], genuine signatures were collected from two sessions, and training signatures were randomly selected from the first session. This process was repeated ten times.
- Types of forgeries available for training: In general, skilled forgeries are not available, and thus, genuine signatures of other people are used as random forgeries. In [5], skilled forgeries were not supplied for training.
- Parameters of decision threshold: Some researchers determine the parameters for the decision threshold using all of the data, including test data, and other researchers determine them using only training data. In the former case, the evaluated result is termed "ideal" [6].

Forgery Data

The measured false match rate, which is useful for predicting the forgery resistance of a signature recognition system, depends on the degree of effort with which signatures of other persons are tried to be imitated. Different types of signature forgeries requiring different skill levels can be distinguished, such as:

- *Random forgery* (also accidental forgery or zero-effort impostor attempt) where an impostor without knowledge of the genuine signature presents any handwriting
- *Simple forgery* where an impostor with knowledge of the genuine signature mimics it from memory without practice
- *Simulated forgery* where an impostor traces a genuine signature without prior practice
- *Skilled forgery* where an impostor mimics the genuine signature after practicing

The quality of skilled forgeries depends on the capability of the forgers, what kind of information they know, how they practice, and how motivated they are.

False match rates measured based on zero-effort impostor attempts, where impostors submit their own biometric characteristics while claiming to be someone else, are meaningful for biometric systems based on physiological characteristics, but are of less relevance for predicting the forgery resistance of signature recognition systems where impostors can easily take action to influence the outcome of recognition attempts in their favor. A skilled forgery is the most difficult to distinguish from a genuine signature. Therefore, for a reliable prediction of the forgery resistance of signature recognition systems, the measurement of false match rates should be based on skilled forgeries. In [5], random forgeries and skilled forgeries were separately used for evaluation.

Signature Databases

For many years, there were no publicly available databases. Therefore, each researcher needed to generate his or her own private databases and evaluate algorithms using them. These databases were not shared with other researchers, and thus, it was difficult to compare the performance of algorithms under the same conditions. This is due to the difficulty in providing forgeries that possess a sufficient level of quality. Forgers, imposters, and seasoned document examiners refuse to or are reluctant to furnish evidence of skillful forgery because such evidence is considered valuable private property.

In the meantime, several on-line signature databases have been developed and made available by several research groups. For example, databases collected by Munich and Perona [7], MCYT [8], BIOMET [9], SVC2004 [5], and MYIDEA [10] are available. These databases contain genuine signatures and skilled forgeries of varying quality. Forgeries in some databases were written by forgers who knew dynamic information about genuine signatures, whereas forgeries in other databases were written by forgers who knew only shape information about genuine signatures. Practical experience with respect to forgeries on bank cheques may be accumulated in places such as Brazil, or in criminal science studies in police institutions [11].

Databases are useful to compare different signature recognition algorithms. However, each database is collected based on a different policy, and they inevitably differ in nature. Thus, evaluation using only a database is not sufficient to determine the overall performance.

Off-line Signatures

Introduction

Off-line signature recognition is the recognition of handwritten signatures based on a two-dimensional gray image obtained by an item of equipment such as a scanner, an optical reader, or a digitizer. Figure 1 depicts an example of an off-line signature.

Since the invention of writing in human society, the signature on a document (or picture, including monochrome brush painting) has been the most common means of authenticating the writer (or painter) of the document (or picture). Not only leaders but also persons accorded with responsibility in various capacities have had to put their signatures on paper and recognize those of the others. Thus, signature recognition has naturally been done off-line. It is only with the recent development of an on-line technology for biometric recognition that the relevance of off-line recognition has been reduced. The relative characteristics of on-line and off-line methods of signature recognition will be discussed later.

Signature Recognition. Figure 1 Off-line signature image.

The written name of the writer was originally used as signature. In the course of the traditional use of signatures, people started including symbols and distorting them in order to increase their uniqueness and beauty. In general, this made it impossible to recover the writers' names from such signatures. Further, some signatures, like names, are merely personal signs that help establish authenticity. In this sense, off-line signature recognition merely entails pattern recognition of questionable images on a two-dimensional space by referring to registered reference images, which may be reduced to a template [3, 12, 2].

The following sections describe the basic modules of the recognition procedure (i.e. data capture, preprocessing, quality assessment, feature extraction, formatting, comparison, decision) for off-line signatures [13].

Basic Operation

Cutting off

In this module, a rectangular space containing the entire image is identified, and the gray image of this area is acquired using an appropriate item of equipment such as a scanner.

Preprocessing

The size of questionable images is not always the same as that of the reference images. Normalizing the size of a questionable image is necessary in order to compare it with templates and thereby facilitate recognition. The background color of the sheet is not always white. In fact, all cheques contain colored patterns for preventing counterfeits, as shown in Fig. 2 [14]. As a result, eliminating the effect of the background pattern and the adjustment of gray levels is inevitable in preprocessing. In general, the binarizing of gray levels up to a certain threshold is preferred to obtain a steady performance of the system.

Normalizing the angle to the baseline (if it exists) proves to be effective in adjusting the signature

Signature Recognition. Figuer 2 An example of background on a traveler's cheque.

position, because the angle (or gradient) is one of the most important pieces of information that help differentiate off-line signature images. In the case of digitized images, rotating an image may add some noise to deteriorate the transformed patterns.

Feature Extraction

There are numerous proposals for the method of feature extraction. In most of them, two types of principles, i.e., pattern matching and multivariate analysis, are combined to construct an effective procedure for similarity measurement and decision-making.

In pattern matching, we place a questionable image on a template for measuring the difference in patterns. The simple matching of gray levels on meshed pixels does not yield robust estimates of similarity due to the shift in the location of images. Therefore, the blurring or smoothing of gray levels on any pixel with respect to adjacent pixels, using a suitable weighted average, is adopted as a method of feature extraction. A set of shadow masking may be considered to apply this principle. The measurement of the fringe of the image is also a technique for feature extraction in pattern matching. The accumulation of black pixels in a binary image along a certain direction such as the x- or y-axis yields a one-dimensional pattern in the form of a time series with the coordinate considered as time.

For a multivariate analysis, multiple variables must be properly defined. The frequency of properly selected primitive patterns can be treated as variables for analysis. The number and location of pikes in the image strokes along a certain direction may serve as variables required to lend features. In certain situations, the gray levels of pixels may act as variables, although the number of variables could go up to 262,144 when the pixels are defined in a mesh of 256 × 1024. The reduction of dimensionality is an important issue to be resolved. The size of the envelope or gradient of strokes also may be used as a feature.

The measurement of frequencies of various local arc patterns on an image yields multiple variables; this is possible by the application of a hybrid principle that combines pattern matching and multivariate analysis.

Measurement of Similarity

For pattern matching, a weighted sum of local similarities may be an index for evaluating the total similarity. When the x- or y-axis is regarded as the time axis, the ▶ dynamic programming comparison method and the framework of a ▶ hidden Markov model may be useful as the tool to construct a good similarity index [15].

In a multivariate analysis, the Euclidian distance, ▶ Mahalanobis distance, Kullback-Leibler divergence, or the deviance based on some probabilistic models can be used, depending on the situation. Linear discriminant functions, quadratic discriminant functions, ▶ support vector machine, or various nonparametric statistics for discrimination are also used for discrimination, although they do not explicitly measure similarity [16, 17, 18].

Judgment on Authenticity

Since there are several criteria of variability among writers, simple judging procedures cannot, in general, extract good performance from the devised system. A multi-step procedure is more practical for achieving better results. Therefore, the technique of fuzzy logic may be effective [19]. The incorporation of the hidden Markov model or neural network technology in the judgment system may also be useful for ensuring reasonably correct judgment.

The critical value or boundary line for judgment should be determined based on the balance of the possibilities of false acceptance and false rejection; here, each possibility is, in general, evaluated by the proportion of the incidence of false acceptance or rejection in the test sample used for designing the system. An increase in false rejections may have grave repercussions on the use of cheques, while an increase in false acceptances may cause serious damage to businesses.

Signature Recognition. Figure 3 Examples of reference images (left), genuine probes (center), and imitated forgeries (right) for a Japanese signature.

Review of Judgment

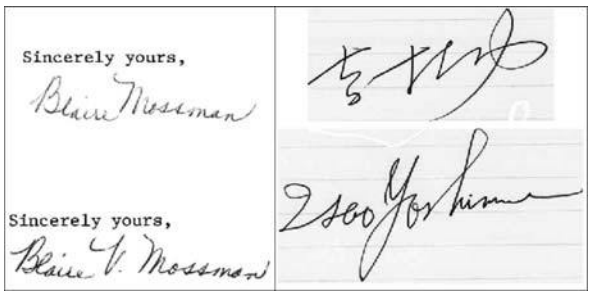
The variability in written images is less than that in the motion of writing, because the writer can adjust his motion by visually tracking the pen movement. This visuality helps seasoned imposters to carry out skilled forgery after extensive training. Even when the imposter is not very skillful, the level of forgery is not very inferior, as seen in Fig. 3. As a result, off-line signature recognition is not a guaranteed way of identifying or verifying identity of an individual. This factor necessitates that judgment be reviewed using other means such as identification cards or passwords, depending on the gravity of the consequences of misjudgment.

Template Construction

A template is provided for measuring the similarity of the questionable image with authentic signature images registered in a database or saved in a system. In the availability of only one authentic image, it should be processed in exactly the same manner as questionable images, in order to construct a template. In the case wherein multiple images for the same signature are stored in a reference database, there are various means of constructing a template. One method is to amalgamate a set of images into an image using a suitable tool. Another method is to select a representative image that is considered as the preferred image.

Influence of Writing System and Nationality

There are many letter systems worldwide, such as Latin, Chinese, Arabic, Cyrillic, Japanese, and Hangul. The distinctiveness of signatures is highly dependent on the letter system. With respect to the manner of writing, letters are written separately in Chinese, Japanese, and Hangul, whereas in English, French, and German, each word is considered as one unit. In European letter systems, a name is considered one word, whereas in some Asian letter systems, it is regarded as a set of disparate characters. The distinctiveness of signatures is not as high in the latter group of countries as it is in the former group due to the social habit of writing [20, 21, 22]. In the Arabic letter system, words are written from right to left.



Signature Recognition. Figure 4 Examples of multiple signatures.

With respect to the use of signatures, a study of the Japanese case reveals singular circumstances. In 1883, the Japanese government legislatively forced all Japanese citizens to use red stamps called *hanko* or *inkan* for official authentication. Further, the Japanese are taught to write their names in the print form in their childhood. As a result, Japanese signatures generally lack uniqueness.

In countries such as Indonesia, the government allows the people to periodically change their signatures for registration in order to maintain security. Given such a situation, the period in which the concerned signature was written is important to verify the authenticity.

Some people have multiple signatures depending on their use, examples of which are shown in Fig. 4.

In the left side of this figure, the middle name is inserted in one signature, while different letter systems produce different signatures for the same name, as seen in the right side of the figure. People in countries that have their own letter system can devise their signatures based on their native alphabet, although they may often use signatures written in Latin letters due to the use of English as the *lingua franca* internationally.

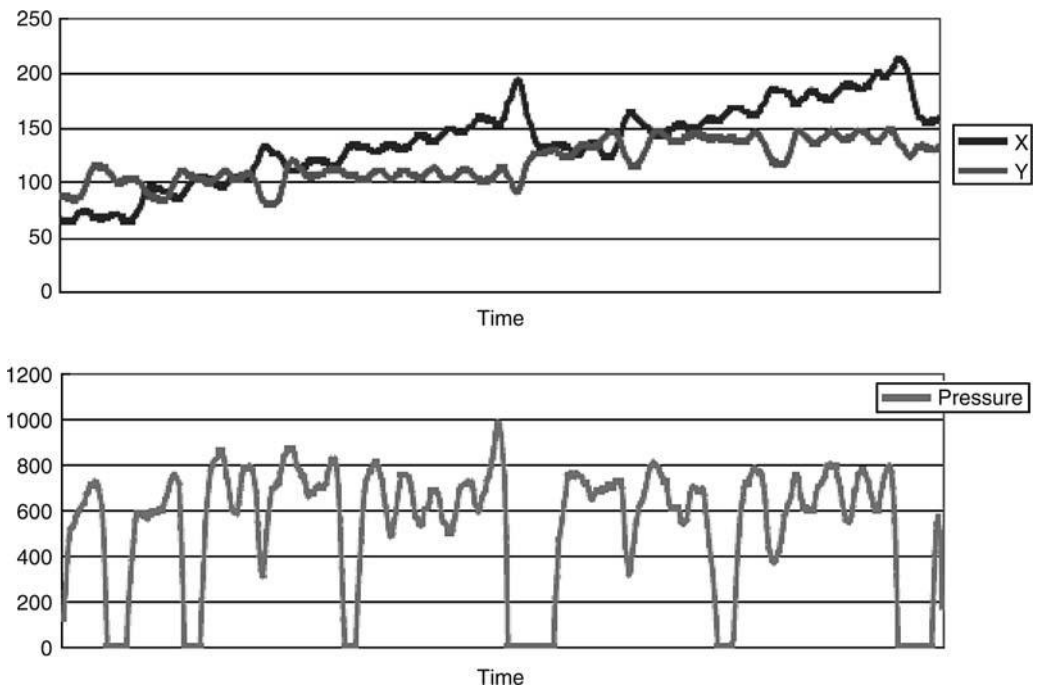
On-line Signatures

Introduction

On-line signature verification uses data obtained while a signature is being written. The data obtained during the process of writing a signature is called an on-line signature. Figure 5 depicts a sample of on-line signature data. On-line signature verification is based on the hypothesis that the writing style of a signature differs from person to person and cannot be easily forged.

On-line signature verification verifies whether an input signature is a genuine signature or a forgery. Ideally, it is a two-class partitioning problem; however, it is not an easy problem to solve, because of the following reasons:

- People do not reproduce their signature exactly each time. Characteristics of the writing manner of genuine writers can change over time. There is, necessarily, intra-class (intra-person) variability. In contrast, forgers attempt to make their forged signatures as similar as possible to genuine signatures, and thus inter-class (inter-person) variability decreases. Therefore, it is difficult to distinguish between genuine signatures and forgeries.
- Both the number and type of signatures available for training are often severely limited. As on-line signature verification is a two-class partitioning problem, general pattern recognition techniques can be applied if enough data is available from both the classes. In practice, however, only a few genuine signatures are available from the genuine class. Moreover, there are several types of forgeries in the forgery class, but only a few types of forgery can be collected for the following reasons: The forgeries that are most similar to genuine signatures and the most difficult to distinguish from genuine signatures will be signatures that were produced by imitating genuine signatures well. Because genuine signatures differ from writer to writer, well-imitated forged signatures should be collected for every writer; however, this is extremely difficult. Scarcity of genuine



Signature Recognition. Figure 5 On-line signature data.

training data exist in all biometric methods; however, scarcity of forgeries exist only in methods that must prepare for imitation attacks.

A variety of algorithms have been proposed for on-line signature verification. Early work carried out up to 1999 is summarized in [2, 3, 23]; subsequently, many studies continue to be reported [6, 7, 24, 25, 26]. A first on-line signature verification competition was held in 2004 [5].

Basic Operation

Data Acquisition

Several data acquisition devices are used for on-line signature verification, for example, digital tablets with pens, pen displays on Tablet PC's, touch screens on PDA's, data acquisition pens, and cameras [6, 7]. These acquisition devices provide ▶ pen coordinate information as time-series data. In addition to **pen coordinate** information, some devices can acquire information about ▶ pen pressure and ▶ pen tilt (▶ altitude, ▶ azimuth) as time-series data.

Preprocessing

After data acquisition, preprocessing such as resampling [6, 7, 27], noise filtering [26], rotation normalization [7], and size normalization [25] is conducted to suppress insignificant information that is expression of random variation, but not of individual signature dynamics, and to even out differences between data captured with different capture devices.

Feature Extraction

This process extracts discriminative features from the acquired data. The features should allow to distinguish a genuine signature from forgeries and be suitable for automatic comparison.

Comparison (Similarity/Dissimilarity Measure)

In the next step, similarity/dissimilarity scores between the extracted features for the input signature and a reference associated with the claimed identity are computed. On-line signature verification algorithms can be classified into parameter-based and function-based approaches, depending on the features compared [3].

In the parameter-based approach, N features, such as average writing speed, total signature duration, and the number of pen ups/pen downs, are extracted as parameters that represent the signature data [26, 28]. In this approach, signature data Sig_{para} after feature extraction can be represented as an N-dimensional parameter vector

$$Sig_{para} = (p_1, p_2, \ldots, p_N), \tag{1}$$

where p_n is the n-th extracted feature, and each feature has a scalar value. In the parameter-based approach, Euclidean distance, weighted Euclidean distance, or correlation are computed as similarity/dissimilarity scores.

In the function-based approach, several features, such as **pen coordinates**, ▶ velocity, and acceleration, are extracted as a sequence of data. In this approach, signature data Sig_{func} after feature extraction can be represented as

$$\begin{aligned} Sig_{func} &= \{sig_{func}(t)\}_{t=1}^{T} \\ &= \{(f_1(t), f_2(t), \ldots, f_N(t))\}_{t=1}^{T}, \end{aligned} \tag{2}$$

where $f_n(t)$ is the n-th feature and T is the number of sample points. Each feature is a function of t. In the case where the features are time-series data, t stands for a time stamp and T is the total duration. In the function-based approach, there are different comparison methods:

- If the reference is also a feature set, two sets of functions are compared, namely, the reference and the features extracted from the input signature. These functions have different durations and are nonlinearly distorted with respect to each other. Then, an elastic comparison algorithm such as dynamic time warping is applied to compute dissimilarity scores [7, 24, 25, 27].
- If the reference is an enrollee-specific function, such as a statistical model, similarity scores are computed using the statistical model and the features extracted from the input signature. For example, the Hidden Markov Model can be used as a statistical model, and probabilities are computed as similarity scores such that the input features are reproduced from the model [26].

After the comparison process, some studies apply score normalization [6, 26]. In the case where multiple scores are computed, score-level fusion strategies [29] are adopted in some studies [6, 7, 25, 26, 27]. In the case where statistical models are used in score-level

fusion [25], model parameters should be estimated in the enrollment process.

Decision Making

In the decision making process, a final decision is made as to whether or not the input signature is considered a genuine signature or a forgery. Some researchers use a classifier for decision making [24], though generally a decision is made by comparing the computed scores with a decision threshold. For example, when similarity scores (a larger score indicates a better match) are used, a decision is made based on

$$\text{Signature} = \begin{cases} \text{Genuine signature if score} > \text{threshold} \\ \text{Forgery if score} \leq \text{threshold} \end{cases} \tag{3}$$

where threshold is a decision threshold. A decision-level fusion strategy [29] can also be used in combination with this decision rule [28].

The decision result can be changed by adjusting the decision threshold value; thus, the decision threshold should be set according to the expected security level. Two types of decision threshold are possible: a writer-dependent threshold (user-dependent threshold) and a common threshold (global threshold). It is reported that a writer-dependent threshold yields better results [6]; however, setting a good user-dependent threshold for each user is a difficult problem. In both types, some parameters estimated in the enrollment process are used to set the decision threshold.

Enrollment

Signatures input during enrollment are called ▶ training signatures. Feature sets are extracted from the **training signatures**, and these extracted feature sets are used in the enrollment process.

For the parameter-based comparison approach, the extracted feature sets are directly enrolled as reference data together with an identifier [28].

For the function-based comparison approach, two types of enrollment are possible depending on the comparison approach adopted, namely, reference-based and model-based [26]. In reference-based enrollment, the feature sets extracted from **training signatures** are directly enrolled as reference data together with the identifier [6, 7, 25, 27]. In model-based enrollment, the extracted features sets are used to estimate statistical models, and these models are enrolled as reference models $f(\cdot;\Theta)$ together with the identifier [26]. Here, $f(\cdot;\Theta)$ is a statistical model and Θ is a parameter set of the statistical model.

Summary

Signature recognition is a historically established method for the authentication of an individual or a document. It is recognized as a common technique worldwide and therefore, is willingly accepted as a means of person authentication by ordinary people. Moreover, signature verification has the notable feature that the signature can be modified in the event that it is compromised. However, since a signature image can be easily copied by other people and there are many people whose signatures are quite variable, the capability of off-line signature recognition as an exclusive means of person authentication is limited.

Many pen input devices, such as tablet PC's and PDA's, are now available that allow to capture the signature dynamics, which is more difficult to copy than signature images. On-line signature verification is thus a promising candidate for person authentication methods. However, the performance of on-line signature verification needs to be improved before it can be used for high-security applications.

Related Entries

▶ Biometrics, Overview
▶ Performance Evaluation

References

1. ISO TC 68/SC 2: Financial services – Biometrics – Security framework. International Standard ISO 19092 (2008)
2. Plamondon, R. ed.: Progress in automatic signature verification. World Scientific, Singapore (1994)
3. Plamondon, R., Lorette, G.: Automatic signature verification and writer identification – the state of the art. Pattern Recognit. **22**, 107–131 (1989)
4. Deng, P., Jaw, L., Wang, J., Tung, C.: Trace copy forgery detection for handwritten signature verification. In: 37th Annual 2003 International Carnahan Conference on Security Technology, pp. 450–455 (2003)
5. Yeung, D., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G.: SVC2004: First international signature verification competition. Number 3072 in Lecture Notes in Computer Science, Springer, Hong Kong, China 16–22 (2004)

6. Jain, A., Griess, F., Connell, S.: On-line signature verification. Pattern Recognit. **35**, 2963–2972 (2002)
7. Munich, M., Perona, P.: Visual identification by signature tracking. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 200–217 (2003)
8. Ortega-Garcia, J., Fiérrez-Aguilar, J., Simon, D., Gonzalez, J., Faundez-Zanuy, M., Espinosa, V., Satue, A., Hernaez, I., Igarza, J.J., Vivaracho, C., Escudero, D., Moro, Q.I.: MCYT baseline corpus: a bimodal biometric database. IEEE Proc. Vis. Image Signal Process. **150**, 395–401 (2003)
9. Garcia-Salicetti, S., Beumier, C., Chollet, G., Dorizzi, B., les Jardins, J., Lunter, J., Ni, Y., Petrovska-Delacrétaz, D.: Biomet: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. Number 2688 in Lecture Notes in Computer Science, Springer, Guildford, UK 845–853 (2003)
10. Dumas, B., Pugin, C., Hennebert, J., Petrovska-Delacrétaz, D., Humm, A., Evéquoz, F., Ingold, R., Rotz, D.: Myidea – multimodal biometrics database, description of acquisition protocols. 59–62 (2005)
11. Sabourin, R.: Off-line signature verification: Recent advances and perspectives. [30] 84–98
12. Leclerc, F., Plamondon, R.: Automatic signature verification and writer verification: The state of the art 1889–1993. Int. J. Pattern Recognit. Artif. Intell. **8**, 643–660 (1993)
13. Impedovo, S., Simon, J., eds.: 1From Pixels to Features III: Frontiers in Handwriting Recognition. North Holland, Amsterdam (1992)
14. Yoshimura, I., Yoshimura, M.: Off-line verification of japanese signatures after elimination of background patterns. Int. J. Pattern Recognit. Artif. Intell. **8**, 53–68 (1994)
15. Yoshimura, I., Yoshimura, M.: An application of the sequential dynamic programming matching method to off-line signature verification. [30] 299–310
16. Justino, E., Bortolozzi, F., Sabourin, R.: A comparison of SVM and HMM classifiers in the off-line signature verification. Pattern Recognit. Lett. **26**, 1377–1385 (2005)
17. Justino, E., Bortolozzi, F., Sabourin, R.: Off-line signature verification using HMM for random, simple and skilled forgeries. 1031–1034 (2001)
18. Kalera, M., Srihari, S., Xu, A.: Off-line signature verification and idenitification using distance statistics. Int. J. Pattern Recognit. Artif. Intell. **18**, 1339–1360 (2004)
19. Hairong, L., Wang, W., Wang, C., Zhuo, Q.: Off-line signature verification and forgery detection using fuzzy modeling. Pattern Recognit. **38**, 341–356 (2005)
20. Hanmandlu, M., Yusof, M., Madasu, V.: Off-line chinese signature verification based on support vector machines. Pattern Recognit. Lett. **26**, 2390–2399 (2005)
21. Ismail, M., Gad, S.: Off-line arabic signature recognition and verification. Pattern Recognit. **33**, 1727–1704 (2000)
22. Yoshimura, I., Yoshimura, M.: Evaluation of signature quality as a function of nationality via an off-line signature verification system. Intell. Automat. Soft Comput. **7**, 195–203 (2001)
23. Plamondon, R., Srihari, S.: Online and off-line handwriting recognition: a comprehensive survey. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 63–84 (2000)
24. Kholmatov, S., Yanikoglu, B.: Identity authentication using improved online signature verification method. Pattern Recognit. Lett. **26**, 2400–2408 (2005)
25. Muramatsu, D., Kondo, M., Sasaki, M., Tachibana, S., Matsumoto, T.: A marvkov chain monte carlo algorithm for bayesian dynamic signature verification. IEEE Trans. Inform. Forensic. Secur. **1**, 22–34 (2006)
26. Fiérrez-Aguilar, J., Ortega-Garcia, J.: On-line signature verification. In Jain, A., Flynn, P., Ross, A. (eds.): Handbook of Biometrics, pp. 189–209. Springer, New York (2008)
27. Nalwa, V.: Automatic on-line signature verification. Proc. IEEE **85**, 215–239 (1997)
28. Lee, L., Berger, T., Aviczer, E.: Reliable on-line human signature verification systems. IEEE Trans. Pattern Anal. Mach. Intell. **18**, 643–647 (1996)
29. Ross, A., Nandakumar, K., Jain, A. (eds.): Handbook of Multibiometrics. Springer, New York (2006)
30. Nabeel, N., Bortolozzi, F. (eds.): Advances in Document Image Analysis. First Brazilian Symposium, BSDIA Number 1339 in Lecture Notes in Computer Science, Springer, Curitiba, Brazil (1997)

Signature Sample Synthesis

Liang Wan[1], Zhouchen Lin[2]
[1]City University of Hong Kong, Hong Kong, China
[2]Microsoft Research Asia, Beijing, China

Synonyms

Signature synthesis; Handwriting synthesis; Handwriting sample synthesis

Definition

Signature sample synthesis is the generation of synthetic signature from a user's signature samples. It is a special case of handwriting sample synthesis which generates novel handwriting in a particular person's handwriting style. A handwriting or signature synthesis system has two basic modules: the modeling module and the synthesis module. In the modeling module, the system collects handwriting/signature samples of a specific writer (online or offline), and identifies and stores the basic characteristics of samples (for example, shape and spatial layout). In the synthesis module,

synthetic ▶ glyphs are generated from the stored templates, and they may be further aligned and connected to form synthetic handwriting data.

Introduction

Like many biometric characteristics, such as face, fingerprint, and iris, signature/handwriting has been widely accepted by people as an effective way to identify a specific writer. Historically, signatures and forged handwritings have always been of interest to forensic experts. Many signature verification technologies have been reported in the literature to detect handwriting forgeries [1, 2]. Signature sample synthesis [3], as an inverse biometrics problem, is the process of generating synthetic signatures that mimic real signature samples. The literature on signature sample synthesis is quite rare. However, it is a special case of handwriting sample synthesis [4, 5], which generates novel artificial handwriting in a person's handwriting style. The article begins with handwriting sample synthesis techniques and then narrows down to signature sample synthesis.

Handwriting sample synthesis has become active in recent years, because the flourish of pen-based devices, such as Tablet PCs, touch-screen mobile phones, personal digital assistants (PDAs), and electronic whiteboards, has brought users more natural communication ways in human–computer interaction. In many situations, writing with a pen on the screen is more convenient than typing on the keyboard. Yet, many users find that keyboards are more efficient than handwriting because typing is faster than writing, and his/her handwriting may become illegible after long-time writing. Handwriting sample synthesis addresses this dilemma by converting ASCII text to handwriting that is close to the user's personal handwriting. For those people who prefer handwriting personal letters, greetings, and compliments, handwriting sample synthesis adds a personal touch to communications. Like wallpapers and favorite software settings, synthesized handwriting also contributes to the personalization of one's computing devices. Moreover, it can always generate legible handwriting and free the user from lengthy and stressful writing, for example, while preparing many hand-written documents such as greeting cards with different content [6].

Handwriting sample synthesis is helpful to build a signature/handwriting recognizer which heavily depends on the size and quality of the training set [7]. It can generate a large database of handwriting/signature samples that look natural. This not only greatly reduces the manual intervention in the preparation of hand-written sample, but also provides the ability to perform operational testing in a laboratory environment. The automated synthesis is also useful to evaluate existing signature/handwriting verification methods [8], including the accuracy and reliability against fraudulent signature/handwriting. In addition, this technique can help forensic examiners [9] to understand the key factors that affect a person's handwriting or signature.

Characteristics of Handwriting/Signature

As a behaviorial biometrics characteristic, handwriting/signature is affected by various factors which the synthesis process should consider. For instance, the signatures are quite different when the specific writer writes in different languages (the article mainly considers the English language, which contains a small set of single characters, and the combination of individual characters in a linear fashion forms various words). Writing with different digital pen devices can also cause direct changes in the appearance of handwriting. Likewise, the person's mood, his/her hand health and the surrounding environment are also possible factors to affect the acquisition process of handwriting samples. Despite those factors, signatures from a single individual tend to be different even using the same digitizing device. For different people, the character shapes can vary greatly and the amount of shape variation may also differ from person to person. In addition, people may tend to write handwriting/signature in a ▶ cursive style or in a partially cursive and partially ▶ handprint style, which makes the problem of synthesis more difficult.

As suggested by handwriting analysis techniques in forensic inspection [10], the specific features that are easily noticeable to ordinary people to distinguish different handwriting styles include: (1) the glyph and the size of single characters; (2) the pressure distribution and the slant of handwriting; (3) the relative sizes of the middle, the upper, and the lower zones of letters; (4) the existence and the shape of head, connecting, and tail parts; (5) the letter, the word, and the line spacings; (6) the embellishment in strokes or character

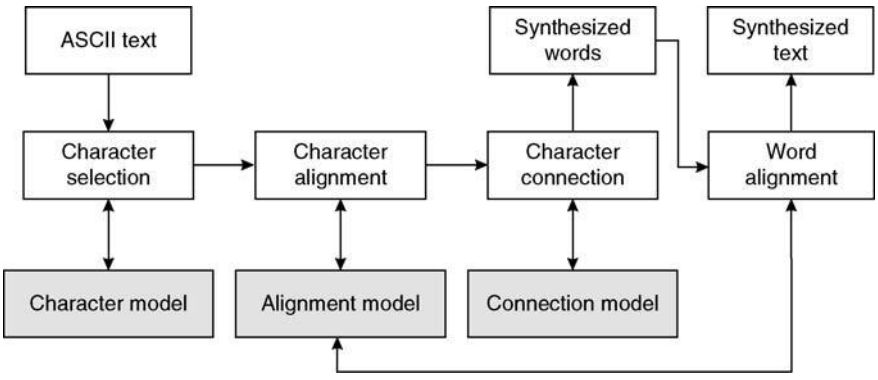
glyphs; and (7) the simplified or neglected strokes. These features can be roughly classified into three types: features of character glyphs, spatial layout of characters, and connection between characters. Different from the English language, the oriental languages, such as Korean and Indian, often contain a large number of characters that share the same small set of strokes, and the characters are usually unconnected. Therefore, the handwriting characteristics of the oriental languages are depicted by two types of features: features of strokes and spatial layout of strokes to form characters [11, 12]. For simplicity, the following discussion focuses on the synthesis of English scripts.

Modeling Process

To generate signatures that look natural, it is important to model the characteristics of signatures/handwriting samples. Figure 1 shows the general outline of the modeling process. It collects handwriting data of a specific user and learns his/her handwriting model. In the modeling process, the handwriting samples are first obtained by acquisition devices. For example, digitizing pen-based devices capture the handwriting sample by a sequence of discrete 2D points. Kinematic information such as pressure and duration of writing can be recorded during the acquisition. Handwriting samples may also be acquired via scanning the off-line sample images, but all kinematic characteristics will be lost.

The general handwriting model consists of three parts: a character model which captures the shapes and variations in single characters; an alignment model which controls the spatial layout of the individual characters that form words; and a connection model which simulates how two characters are connected together. The simplest form to represent the signature/handwriting is a planar curve. Hence, the character model often extracts a set of control points to represent the hand-written character glyph [1, 3, 13–17]. Bezier curves or polynomials are then used for curve approximation. The shape variation such as scale, position, and slant can be learnt from multiple samples of one character. Besides the geometric information, physically plausible models have been proposed to model the speed and acceleration in the writing [4, 5]. These kinematic-based models are capable of representing, compressing and reconstructing input handwriting data, but they do not target at synthesizing new handwriting. In contrast, geometry-based models can generate handwriting with natural shape variations and support different handwriting styles, i.e., from handprint style to fully cursive style. The alignment model records the horizontal letter spacings and the relative vertical positions of characters with respect to a horizontal baseline [17]. The connection model may record which character pairs are likely to connect to each other. It may also extract the distribution of concatenation strokes which are formed by the tail and the head parts of adjacent characters [14].

In the modeling process, the system usually requires the users to provide adequate handwriting samples, so a practical concern is to keep user involvement at a reasonable level. In fact, the burden of user involvement in the sample collection process depends upon the handwriting models that the systems use. For example, users are asked to write more than one thousand letter groups in [13]. Besides writing 80–200 words, the work in [14, 15] may need user interaction



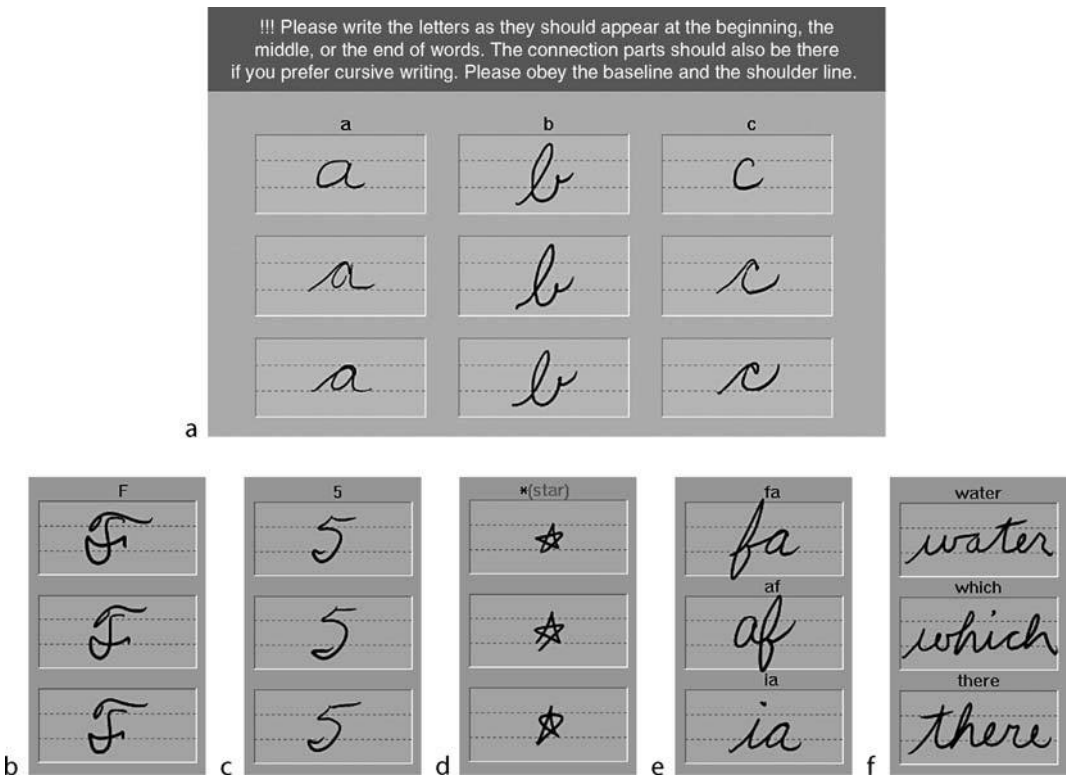
Signature Sample Synthesis. **Figure 1** Block diagram of the model training process.

in order to get good segmentation results. In contrast, the user is only required to input each single character three times, several special pairs of letters and several multiletter words (Fig. 2) in [17].

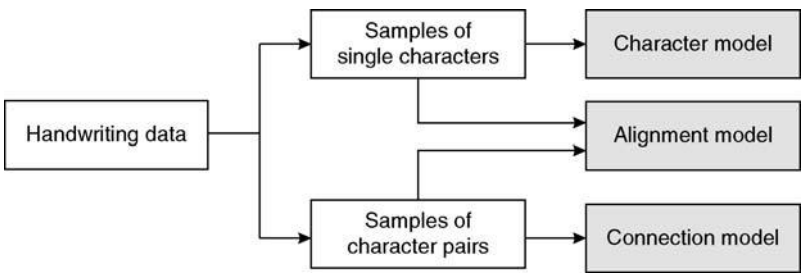
Synthesis Process

The synthesis process should have synthetic data visually similar to the samples and incorporate sufficient variability in synthetic data. This objective, however, is not easily achieved, especially when considering cursive writing styles. Figure 3 shows the basic flowchart of handwriting synthesis systems. For an input ASCII text, each individual character glyph is first generated from the character model. Then the glyphs are arranged and adjacent characters are connected when needed to form a cursive word. The words are further aligned into lines and paragraphs. In the following section, each step of synthesis is detailed.

Character selection generates glyphs based on the stored character models. In [4], the handwriting



Signature Sample Synthesis. Figure 2 The user interface in [17] to collect user handwriting samples: samples of lowercase letters, capital letters, digits, punctuations, special letter pairs, and multi-letter words.



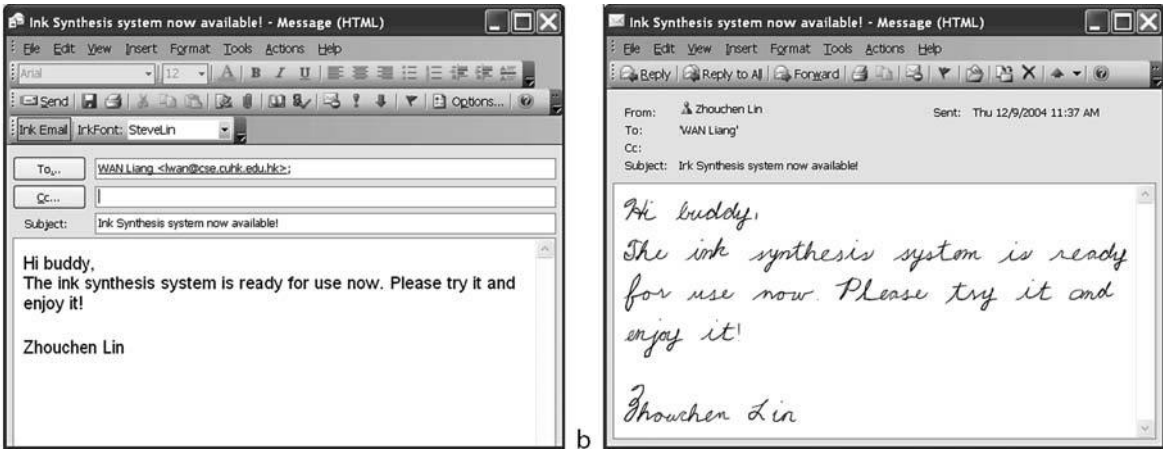
Signature Sample Synthesis. Figure 3 Block diagram of the handwriting sample synthesis process.

sample is modeled by the writing velocity equations. Then a similar synthetic data are generated by setting initial conditions to the equations. In [14, 15], the character model is the statistical distribution of control points that are learnt from multiple glyph samples for each character. By randomly sampling the distribution, new handwriting trajectory can be generated. In [17], three instances of characters are stored as the templates which are supposed to appear at the beginning, the middle, and the end of words, respectively. Then according to the character position in the word, one of the three templates is chosen as the initial glyph. After selection, affine transformation (scale, rotation and slant) is applied to the character glyphs. To mimic the variability in the natural handwriting/signature, some randomness is often added in the transformation.

Character alignment then places the glyphs with respect to the baseline, both vertically and horizontally. Vertical alignment is necessary for the generation of smooth handwriting. For example, middle-zone letters ("a," "c," "e," etc.), ascendent letters ("b," "d," etc.), and digits are expected to have their glyph bottoms meet the baseline. Horizontal alignment, on the other hand, separates the bodies of adjacent letters at a distance along the horizontal baseline. However, the adjacent character glyphs may have severe overlapping when their head or the tail parts are too long. As a result, the synthesized handwriting may look weird or it may be hard to produce smooth connection part. Hence, redundant portions of heads/tails may be trimmed to alleviate the overlapping problem [17].

Character connection is developed to simulate cursive handwriting, in which adjacent letters are connected by smooth ligature parts (the head and tail parts of adjacent characters). In [14], a statistical ligature generation model is learnt from handwriting samples. when synthesis, a ligature stroke is generated for every pair of adjacent letters. The final ligature part is determined by jointly deforming the letter strokes and the ligature stroke. This method needs sufficient samples for training. In [15], a delta log-normal model is employed to represent the head/tail parts as pieces of arcs. By changing the arc parameters, the trajectories of letters are deformed to create a smooth ligature part. This model may be interfered by too long overlaps between head/tail parts. The work [17] adopts a high-order polynomial to fit the ligature part. The problem becomes to determine the control points given the head/tail parts of adjacent glyphs. To solve the fitting problem, three constraints are imposed on the ligature: similarity to the original ligature, deformation energy from the original ligature, and smoothness of the ligature. The control points should minimize the sum of these three energy terms.

After the processes mentioned above, an ASCII word is converted to a hand-written word. Then it is natural to synthesize handwriting text by rendering multiple words one by one. Figure 4 shows an example of the communication via emails with handwriting synthesis. The sender types in a text email. The handwriting synthesizer automatically converts it into a handwriting email and sends it to the receiver. Then



Signature Sample Synthesis. **Figure 4** Integration of handwriting sample synthesis system with Microsoft® Office Outlook®.

the receiver finally reads the letter in the sender's personal handwriting style.

Since signature is a special type of handwriting, handwriting synthesis techniques can be adapted for signature sample synthesis. Instead of representing each character in handwriting synthesis, the whole signature can be taken as a single glyph. Then the character modeling is applied to the signature glyphs, for example, extracting the control points or estimating the writing speed. During synthesis, the process of character selection is applied to the learnt models of signature glyphs. By this way, a large database of synthetic signatures can be obtained to test the robustness of an existing handwriting/signature verification or recognition method.

Summary

As an inverse biometrics problem, signature/handwriting sample synthesis has been studied in recent years. Existing synthesis systems can help common users to produce personal signature/handwriting with pleasing visual quality. However, they do not capture all aspects of the handwriting style. For example, the handwriting of people may evolve gradually with their ages. It may add more liveness to the handwriting if the effect of time is considered. Furthermore, the research on synthesis in languages other than English needs more investigation before it becomes accessible to people in different countries.

Related Entries

- ▶ Handwriting Structure
- ▶ Sample Synthesis from Templates
- ▶ Signature Features
- ▶ Signature Recognition

References

1. Singer, Y., Tishby, N.: Dynamic encoding of cursive handwriting. Biometric Cybernetics **71**(3), 227–237 (1994)
2. Gupta, G., McCabe, A.: A review of dynamic handwritten signature verification. Tech. rep. (1997)
3. Oliveira, C.D., Kaestner, C., Bortolozzi, F., Sabourin, R.: Generation of signatures by deformations In: BSDIA '97: Proceedings of the First Brazilian Symposium on Advances in Document Image Analysis, pp. 283–298 Springer, London, UK (1997)
4. Plamondon, R., Maarse, F.: An evaluation of motor models of handwriting. IEEE Trans. Pattern Analy. Machine Intell. **19**(5), 1060–1072 (1989)
5. Li, X., Parizeau, M., Plamondon, R.: Segmentation and reconstruction of on-line handwritten scripts. Pattern Recogn. **31**(6), 675–684 (1998)
6. FontGod Corperation: Handwriting fonts. http://www.fontgod.com/ (2007)
7. Zheng, Y., Doermann, D.: Handwriting matching and its application to handwriting synthesis. In: Proceedings of Eighth International Conference on Document Analysis and Recognition, vol. 2, pp. 861–865 (2005)
8. Brault, J.J., Plamondon, R.: A complexity measure of handwritten curves: modelling of dynamic signature forgery. IEEE Trans. Syst. Man Cybernetics **23**, 400–413 (1993)
9. Tappert, C.: Handwriting synthesis of a particular writer's style. http://csis.pace.edu/ctappert/research/researchtopics.htm (2007)
10. Srihari, S., Cha, S., Arora, H., Lee, S.: Individuality of handwriting. J. Forensic Sci. **47**(4), 856–872 (2002)
11. Lee, D.H., Cho, H.G.: A new synthesizing method for handwriting Korean scripts. Int. J. Pattern Recogn. Artif. Intelligence **12**(1), 46–61 (1998)
12. Jawahar, C., Balasubramanian, A.: Synthesis of online handwriting in Indian languages. In: Proceedings of International Workshop on Frontiers in Handwriting Recongnition. La Baule, France (2006)
13. Guyon, I.: Handwriting synthesis from handwritten glyphs. In: Proceedings of Fifth International Workshop on Frontiers of Handwriting Recognition, pp. 309–312 Colchester, England (1996)
14. Wang, J., Wu, C., Xu, Y.Q., Shum, H.Y., Ji, L.: Learning based cursive handwriting synthesis. In: Proceedings of Eighth International Workshop on Frontiers of Handwriting Recognition, pp. 157–162. Ontario, Canada (2002)
15. Wang, J., Wu, C., Xu, Y.Q., Shum, H.Y.: Combining shape and physical models for on-line cursive handwriting synthesis. Int. J. Doc. Analy. Recogn. **7**(4), 219–227 (2005)
16. Nikitin, A.V., Popel, D.V.: Signmine algorithm for conditioning and analysis of human handwriting. In: Proceedings of the International Workshop on Biometric Technologies, pp. 179–190 Calgary, Alberta (2004)
17. Lin, Z., Wan, L.: Style-preserving English Handwriting Synthesis. Pattern Recogn. **40**, 2097–2109 (2007)

Signature Similarity Computation

▶ Signature Matching

Signature Synthesis

▶ Signature Sample Synthesis

Signature/Sign Recognition

▶ Signature Recognition

Silhouette

The set of pixels in an image corresponding to an object of interest. In gait recognition, the object of interest is the human subject to be recognized. These are outlines of persons or object of interest in images or videos. In gait recognition, these are represented as 2D (filled) shapes of walking persons. Most common usage includes both the inside and the boundary of the shape to be part of the silhouette. Silhouettes are typically detected by subtracting the background image from any given image, followed by some simple hole filling and clean up operations.

▶ Gait Recognition, Silhouette-Based

Silhouette Analysis for Gait Recognition

▶ Gait Recognition, Motion Analysis for

Similarity Metric

Any of a large number of scalars assessing the degree of similarity, or of dissimilarity, between objects such as biometric features or templates is similarity metric. Examples include: correlation statistics, vector projections (cosine or inner products), and fractional Hamming Distance (the fraction of bits that differ between two bit sequences of equal length). Similarity metrics usually obey the axioms of distances (symmetry; non-negativity; and triangle rule inequality), but they can incorporate nonlinear or non-Euclidean topologies (e.g., city-block distances between points in a space versus straight-line distances). A similarity metric between biometric templates, after being suitably normalized, is usually the input into a decision process that renders a judgment about whether they should be classified as "same" or "different."

▶ Palmprint Matching
▶ Score Normalization Rules in Iris Recognition

Simplifying Passenger Travel Program

Synonyms and Acronyms

ABG; CANPASS; IRIS; NEXUS; Privium; SPT

Definition

Organized by IATA (the international air transport association representing 230 airlines), the simplifying passenger travel initiative refers to a combination of programs agreed by airlines, airports, and governmental regulatory bodies for the purpose of simplifying the experience of air travel and making it more efficient. Not focused solely on expediting security procedures for departing passengers, SPT programs also include the use of biometrics for automated immigration clearance for arriving passengers, streamlined border-crossing, automated check-in, and enhancing other airport processes. National programs that use iris recognition for these purposes include IRIS (UK): Iris recognition immigration system; Privium (NL); ABG (Germany); and the Canadian and US border-crossing programs, CANPASS and NEXUS.

▶ Score Normalization Rules in Iris Recognition

Simultaneous Capture of Iris and Retina for Recognition

David Usher, Yasunari Tosa, Marc Friedman
Retica Systems, Inc., Waltham, MA, USA

Synonyms

Iris retina biometric fusion

Definition

The simultaneous imaging of the iris and retina for the purposes of biometric recognition. A specialized optical engine is used to illuminate and image both biological features. Acquisition software extracts optimal images from acquired video sequences. The resultant digital images are analyzed and biometric information is extracted and stored. Information from both the iris and retina is combined or fused for the purposes of verifying or determining the identity of the individual. The fixed anatomical proximity of the iris and retina facilitate their simultaneous capture by a single device. Combining the iris and retina traits aims to offer an advantage over unimodal iris and retina biometric systems.

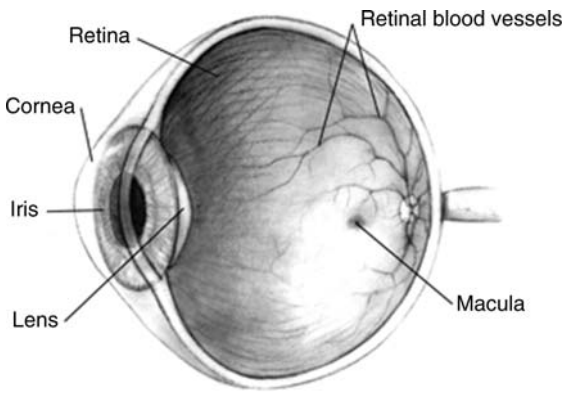
Introduction

Within the field of biometrics, the iris has received considerable attention. Iris biometric systems are commercially available and research is expanding. Only recently has the technology begun to meet its commercial potential. The reasons for this involve meeting the many challenges necessary to create systems that combine high quality imaging, good human factors engineering, and high quality software algorithms for image capture, segmentation, encoding, and matching. Retina-based identification has long been perceived as a robust biometric solution, but very few practical applications or commercially viable products have been demonstrated. EyeDentify Inc. developed a retinal biometric product [1, 2] that demonstrated reasonable performance [3]. However, it suffered from a perception that its human interface was intrusive. Optibrand Ltd. developed a retinal biometric device for the livestock market [4]. More recently, Retica Systems Inc. has developed a biometric acquisition system that combines the retina with iris biometrics [5]. As biometric systems attempt to meet the demands of real world applications, multibiometric systems are receiving considerable attention. It is believed that some of the limitations imposed by unimodal systems can be overcome by using multiple biometric modalities [6]. Both iris and retina systems can suffer from problems associated with unimodal systems. These include noisy data, intraclass variation, nonuniversality, and susceptibility to spoof attacks. Although iris technologies have demonstrated high levels of performance, research with the aim of mitigating poor iris image quality is ongoing, and attempts have been made to improve performance by combining the iris with other biometric traits including fingerprint [7] and face [8]. The fixed anatomical proximity of the iris and retina facilitates their simultaneous capture using a single system. Biometric traits are best combined when their discriminating power is evenly balanced and their content is independent. The topology of the retinal blood vessels is independent of the texture of the iris. It therefore may be possible to improve biometric performance by combining balanced iris and retina recognition technologies into a single device.

Anatomical Background

The eye can be divided into the *anterior* and *posterior* segments, Fig. 1. The iris is found in the anterior segment that also includes the cornea and lens. The iris is constructed of pigmented fibrovascular tissue layered onto a back surface of pigmented epithelium cells. Crypts and freckles add to the observed pattern. The texture variation across the iris is distinctive and it is this information that is encoded forming the iris biometric signal. The retina is found in the posterior segment that comprises the back two-thirds of the eye. Light is refracted by the cornea and lens through the pupil onto the retina, a thin layer of neural cells that lines the interior surface of the eye (the fundus). The fundus, as seen using a digital fundus camera, is shown in Fig. 2 where the images of the left eyes of two identical twins are shown. The retinal blood vessel pattern is the subject of biometric encoding methods. Two major blood vessel systems supply the retina. The outer retinal layers are supplied by a

choroidal blood vessel network. The choroidal vessels form a grid-like pattern and are not generally visible using standard digital fundus cameras and refractive ophthalmoscopes. The inner layers of the retina are supplied by the ► central retinal artery. There is also one main collecting trunk, the ► central retinal vein. These two blood vessels form ► bifurcations as they emerge from the optic disc and branch out through the nerve fiber and ganglion cell layers forming an extended network throughout the retina. The optic disc is the point where the optic nerve breaks out into the retina and can be seen as the bright spots in Fig. 2. Several studies have concluded that the branching patterns of the retinal arterial and venous systems have characteristics of a fractal [9, 10]. It has been suggested that a nonequilibrium Laplacian process could be involved in ► retinal angiogenesis [10] and that fluctuations in the distribution of embryonic cell-free spaces provide the randomness needed for fractal behavior and for the uniqueness of each individual's retinal vascular pattern [9]. This fractal-like growth occurs in the embryonic stages of humans and provides for uniqueness even in the case of identical twins, Fig. 2. The anatomical stability of the iris and retina biometric traits must be considered. While it is expected that both suffer from minimal normal age-related changes, both are effected by various disease states. Large-scale studies specifically addressing the stability of the iris, especially for biometrics, have yet to be performed [11]. This is also the case for retinal biometrics.



Simultaneous Capture of Iris and Retina for Recognition. **Figure 1** Schematic diagram of the human eye (right). (Courtesy of National Eye Institute, National Institutes of Health).

Challenges

The principal challenges for an ocular imaging biometric device broadly fall into three categories encompassing the imaging system, its human interface, and software analysis algorithms:

1. *Imaging system.* The challenge for the imaging system is to record stable images that best illuminate biometric features. In the case of the iris, standards

Simultaneous Capture of Iris and Retina for Recognition. **Figure 2** Fundus camera images corresponding to the left eyes of two identical twins. Retinal blood vessels can be seen to form a branching network centered on the optic disc. The optic disc can be seen as a bright spot near the center of each image. Comparing the two images, the topology of the blood vessels is notably different. (Courtesy of Prof. Michael Larsen, Glostrup Hospital, University of Copenhagen).

[12] dictate two basic requirements: (1) near-infrared illumination must be used, (2) pixel and spatial resolution limits must be met. Illumination across the eye must produce an image with even levels of contrast throughout and with clearly defined iris boundaries. Iris texture should be emphasized and the optical system must minimize obscuring reflections including ► Purkinje images and those from ambient light sources. In the case of the retina, the interior surface of the eye must be imaged through the refracting surfaces of the cornea and lens and through its natural aperture, the pupil. A suitable field-of-view containing a high level of blood vessel detail is required. The blood vessels must show a suitable level of contrast. Optical appliances such as glasses, contact lens and other types of face guards, and masks add additional imaging challenges for both iris and retina systems, as they can contain scratches and generally exhibit poor transmission properties. In addition, in the case of a dual iris-retina system, iris and retina imaging systems should complement each other and steps must be taken such that there is no interference between the two.

2. *Human interface.* Human interface challenges encompass fixation and targeting, illumination considerations, and distance requirements, and are highly dependent on the application of the technology. In general, a biometric device should inconvenience the user as little as possible while facilitating repeatable and stable imaging of the biometric traits. The challenges associated with imaging the retina dictate that, at least in the short term, a passive imaging system with a level of active participation by the user is required. If a passive system is employed, then a suitable alignment tool is required. This fixture must be straightforward and intuitive to use and be capable of aligning the user to a defined degree of accuracy.
3. *Software analysis.* Software analysis challenges involve live acquisition, feature extraction, encoding, and matching. The task of the image acquisition step is to identify from video sequences if any acceptable views of the biometrics have been presented and, if so, to extract and record the best examples. In this context the term *best* means an image from which the biometric signal can be encoded with the highest degree of accuracy. Image acquisition methods must process continuous video sources. Processing constraints are therefore high. A definition of image quality is needed and thresholds must be applied to exclude unacceptable images. There is an inherent trade-off between the image quality thresholds applied during image acquisition, the human interface, and the subsequent efficacy of the encoding. Image quality constraints set too high may result in a more prolonged and difficult user experience or ubiquitous failures-to-acquire. Image quality constraints set too low may compromise encoding and therefore potentially degrade matching performance. Feature extraction techniques are used to reduce the acquired images into biometric signals. Methods must accurately extract the unique features present in the image and efficiently encode them to facilitate matching. For iris analysis systems, the iris must first be located and separated from the rest of the image. This necessarily includes identifying areas of iris occlusion. In the case of the retina, the blood vessel network must be separated from other features within the retinal images. Encoding methods must provide an efficient characterization of the biometric features that facilitates accurate and rapid matching methods. Finally, matching algorithms must define a similarity score such that scores from pairs of signals from the same individual show a high separation from scores generated from signals of different individuals. Operational constraints often dictate that matching must be rapid.

Image Quality Considerations

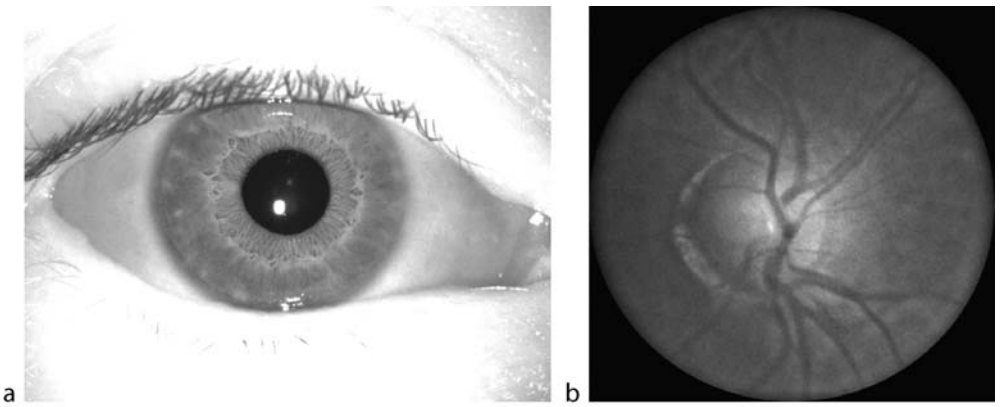
Systems that combine a good image capture hardware design with a straight forward human interface can still suffer from large variations in image quality. Inherent anatomical, behavioral, and environmental variations introduce confounding factors. A large range in iris reflectance can result in low contrast for one of its boundaries (e.g., dark irises can exhibit poor contrast with the pupil. Highly reflective irises can result in a low level of contrast for the iris–sclera boundary). Variations in the reflectivity of the retina can result in a range of retinal image brightness and contrast. Variations in the optical efficiency of the eye and the variable use of glasses or contact lenses can result in a range of achievable focus for a fixed focused optical system imaging the retina, and add unwanted reflections to iris images. Ambient light sources also affect

the size of a user's pupil. In addition to altering the appearance of the iris, pupil scale changes can dynamically affect the illumination on the retina. Variations in user alignment can introduce eye-gaze to iris images that can adversely affect segmentation and can introduce nonlinear transformations to the blood vessel pattern as projected onto a two dimensional retinal image. Standards for iris image quality have been defined [12]. No such standard exists for retinal images.

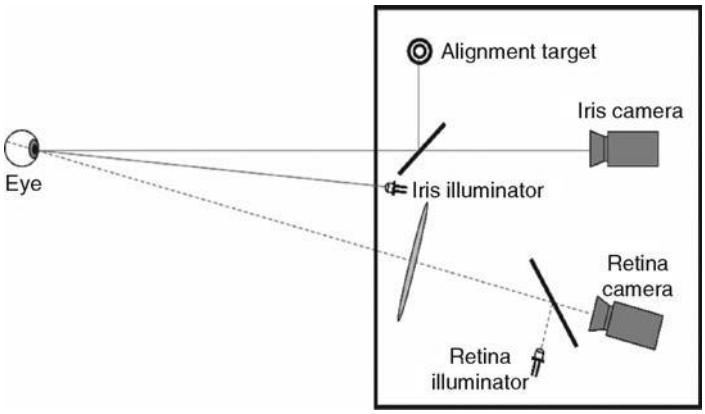
Simultaneous Capture of the Iris and Retina

This section discusses Retica's fusion of its iris and retina technologies for simultaneous capture by a single-presentation biometric device. Solutions to the challenges discussed earlier combine to form a novel iris–retina imaging system, its associated human interface, and proprietary software analysis algorithms for automated acquisition, encoding, and matching. A more detailed description of Retica's dual iris–retina technology can be found in [5].

1. *Imaging system.* Proprietary iris and retina optical systems are arranged in a single housing. Both systems employ standard Video Graphics Array (VGA) resolution cameras. In the iris optical system, illumination is provided by a set of narrowband near-infrared (NIR) LED's. A fixed focus system with a focal length of 27 cm is used. Care was taken to meet the iris image quality standards [12] for resolution, contrast, and noise levels. The resolution of the iris imaging system is 25 pixels per mm. An example of an iris image recorded using Retica's imaging system is shown in Fig. 3(a). The retinal optical system was simplified by using a 10° field-of-view. This relatively narrow field allowed image capture from a greater distance. The retina camera was set at a horizontal angle of 15.5° and 1.5° below line of sight, (see Fig. 4). At this angle its field-of-view was centered on the optic disc region. This region was chosen, as it has a high concentration of blood vessels. In addition, the optic disc is close to the pivot point for eye rotation. It is therefore the most stable part of the retina in terms of transformations in the recorded images as a result of eye movements. Illumination is provided by a narrowband NIR LED reflected off a beamsplitter. An aperture is imaged by a large lens to a 2 mm spot just before the cornea uniformly illuminating the optic disc region. The resolution of the retina imaging system is 90 pixels per mm. Optimal positioning is approximately 27 cm from the front panel. Figure 3(b) shows an example of a captured image of the retina. The optic disc can be seen near the center of the image along with radiating blood vessels.
2. *Human interface.* A beam splitter placed in front of the iris camera (Fig. 4) is used to create a targeting system for the user consisting of a colored illuminating ring surrounding a colored disk containing



Simultaneous Capture of Iris and Retina for Recognition. Figure 3 (a) An example of an iris image captured using Retica's iris optical system and associated software analysis tools. The field-of-view comprises 26 by 20 mm encompassing the whole visible section of the eye. (b) An example of a retinal image captured using Retica's retinal optical system and associated software analysis tools. A 10° field-of-view was centered approximately on the optic disc).



Simultaneous Capture of Iris and Retina for Recognition. **Figure 4** Outline of Retica's dual retina-iris optical system.

crosshairs. When the subject is able to see the disk centered within the ring, they are in correct lateral alignment. The user is then asked to move forward maintaining lateral alignment. In doing so they move through the fixed focus of the iris system and across the optical axis of the retina system. Empirical evidence showed that lateral alignment could be achieved easily even for inexperienced users. However, there was a large range in the expectation of the distance positioning that was required. Visual and auditory cues are used to guide the user as they move towards the device. An optical rangefinder is used to set the color of the targeting system's colored disk to green when the user is at approximately the correct distance. In addition to this, images meeting predefined quality thresholds (see the following section) trigger an audible tone. The user is instructed to continue moving towards the device until the tone stops, then to move slowly backwards again moving through the region coinciding with the tone once more. These steps are repeated until the acquisition process is interrupted when recorded data have met predefined thresholds for quality and quantity. The user must move through the optimal alignment position; they are not required to hold a fixed position. Moderately experienced users could be acquired in less than 1 second. Most inexperienced users were acquired in less than 15 seconds. The active cooperation of the user combined with an intuitive alignment tool reduce the effect of eye gaze in the acquired iris images.

3. *Software analysis.* Iris and retina acquisition algorithms automate the acquisition process. Image quality thresholds for iris and retina image focus are defined along with measures of iris occlusion and retinal blood vessel content. A video rate assessment of image quality is then used to select *best* iris and retina frames as the user passes though optimal alignment. Acquired images are passed on to proprietary iris and retina encoding algorithms. Encoded templates are then matched against existing databases.

Conclusion

The section introduced the concept of simultaneous capture of iris and retina for biometric recognition. Some of the challenges associated with iris and retina biometric systems were discussed. A bimodal system that demonstrated that the iris and retina can be acquired simultaneously was outlined and represents a unique contribution to the field of biometrics. Although unimodal results are good, intraclass variations can present problems for both biometrics. As discussed earlier, anatomical, behavioral, and environmental factors can result in a range in image quality for both the iris and retina. The level of success of a recognition system can be largely defined by how well its encoding and matching methods are able to manage these variations. For example, inexact iris localization because of poor iris image quality or failures in blood vessel segmentation because of poor retinal image quality ultimately result in poor genuine match scores. Commercially available iris recognition systems exhibit nonzero error rates. Problems such as occlusion, motion blur, specular reflections, and pose contribute to

intraclass variations. It may be possible to enhance performance by using a dual iris–retina system. The topology of the retinal blood vessel pattern is completely uncorrelated to the texture patterns on the iris. They therefore represent complementary sources of information. While it is true that they share some ubiquitous failures, (e.g., someone with no eye has neither an iris nor a retina) various obfuscatory factors affecting the iris or the retina are either uncorrelated or anticorrelated. For example, eyelid and eyelash occlusion has no relation to retinal blood vessel detail. Highly dilated pupils that can cause problems for iris systems aid imaging of the retina. Both traits represent strong biometrics potentially facilitating a more balanced fusion than the combination of a strong biometric with a weaker one. Both biometric features are enclosed organs and cannot be altered without endangering vision. However, there is a risk of spoof iris attacks, and antispoofing measures are being actively investigated. Adding a requirement for retinal identification significantly increases the challenge of hoax enrollment. Next generation iris acquisition systems aim to relax constraints imposed on users in terms of capture volume, standoff distance, and motion. While the system proposed by Retica demonstrated straightforward retina acquisition at 27 cm, less-constrained retina acquisition presents a significant challenge. Dual iris–retina capture systems are therefore likely to be restricted to applications that require the highest accuracy, for cooperative users, with relatively constraining human interfaces.

Related Entries

► Eye Features and Anatomy
► Iris Acquisition Device
► Iris Image Data Interchange Formats, Standardization
► Retina Recognition
► Vein and Vascular Recognition, Overview

References

1. Johnson, J.C., Hill, R.B.: Eye fundus optical scanner system and method. US Patent No. 5668842 (1990)
2. Hill, R.B.: Retina identification. In: A.K. Jain, R. Bolle, S. Pankanti (eds.) Biometrics Personal Identification in Networked Society., chap. 6, pp. 123–142. Springer, London (1999)
3. Holmes, J.P., Wright, L.J., Maxwell, R.L.: A performance evaluation of biometric identification devices. Tech. Rep. SANDIA91–0276, Sandia National Laboratory (1991)
4. Golden, B.L., Rollin, B.E., Switzer, R.V., Comstock, C.R.: Retinal vasculature image acquisition apparatus and method. US Patent No. 6766041 B2 (2004)
5. Usher, D., Tosa, Y., Friedman, M.: Ocular biometrics: Simultaneous capture and analysis of the retina and iris. In: N. Ratha, V. Govindaraju (eds.) Advances in Biometrics: Sensors, Algorithms and Systems, chap. 8, pp. 133–155. Springer, London (2008)
6. Ross, A., Jain, A.J.: Multimodal biometrics: An overview. In: Proceedings of 12th European Signal Processing Conference (EUSIPCO), pp. 1221–1224. Vienna, Austria (2004)
7. Nandakumar, K., Chen, Y., Dass, S.C., Jain, A.K.: Quality-based score level fusion in multibiometric systems. In: Proceedings of International Conference on Pattern Recognition (ICPR), pp. 20–24. Hong Kong (2006)
8. Wang, Y., Tan, T., Jain, A.K.: Combining face and iris biometrics for identity verification. In: Proceedings of fourth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 9–11. Guildford, UK (2003)
9. Mainster, M.A.: The fractal properties of retinal vessels: Embryological and clinical implications. Eye **4**(1), 235–241 (1990)
10. Masters, B.R.: Fractal analysis of the vascular tree in the human retina. Annu. Rev. Biomed. Eng. **6**, 427–452 (2004)
11. Wildes, R.P.: Iris recognition. In: A.K. Jain, D. Maltoni, D. Maio (eds.) Biometric systems. Technology, design and performance evaluation, pp. 63–95. Springer, London (2004)
12. ANSI/INCITS: American National Standard For Information Technology. Iris image interchange format. ANSI INCITS 379–2004 (2004)

Skilled Forgery

In signature verification, skilled forgeries represent the imitations that are intentionally performed, that is, where the forger actively tries to imitate the signature from another user.

► Signature Databases and Evaluation

Skin Classification

► Skin Detection

Skin Color Detection

► Skin Detection

Skin Detection

Ahmed Elgammal, Crystal Muang, Dunxu Hu
Department of Computer Science, Rutgers University, Piscataway, NJ, USA

Synonyms

Skin classification; Skin color detection

Definition

Skin detection is the process of finding skin-colored pixels and regions in an image or a video. This process is typically used as a preprocessing step to find regions that potentially have human faces and limbs in images. Several computer vision approaches have been developed for skin detection. A skin detector typically transforms a given pixel into an appropriate color space and then use a skin classifier to label the pixel whether it is a skin or a nonskin pixel. A skin classifier defines a decision boundary of the skin color class in the color space based on a training database of skin-colored pixels.

Introduction

Skin color and textures are important cues that people use consciously or unconsciously to infer variety of culture-related aspects about each other. Skin color and texture can be an indication of race, health, age, wealth, beauty, etc. [3]. However, such interpretations vary across cultures and across the history. In images and videos, skin color is an indication of the existence of humans in such media. Therefore, in the last two decades extensive research has focused on skin detection in images. Skin detection means detecting image pixels and regions that contain skin-tone color. Most of the research in this area has focused on detecting skin pixels and regions based on their color. Very few approaches attempt to also use texture information to classify skin pixels.

As will be described shortly, detecting skin pixels is rather a computationally easy task and can be done very efficiently, a feature that encourages the use of skin detection in many video analysis applications. For example, in one of the early applications, detecting skin-colored regions was used to identify nude pictures on the internet for the sake of content filtering [6]. In another early application, skin detection was used to detect anchors in TV news videos for the sake of video automatic annotation, archival, and retrieval [1]. In such an application, it is typical that the face and the hands of the anchor person are the largest skin-tone colored region in a given frame since, typically, news programs are shot in indoor controlled environments with man-made background materials that hardly contain skin-colored objects. In many similar applications, where the background is controlled or unlikely to contain skin-colored regions, detecting skin-colored pixels can be a very efficient cue to find human faces and hands in images. An example in the context of biometrics is detecting faces for face recognition in an controlled environment.

Detecting skin-colored pixels, although seems a straightforward easy task, has proved quite challenging for many reasons. The appearance of skin in an image depends on the illumination conditions (illumination geometry and color) where the image was captured. Humans are very good at identifying object colors in a wide range of illuminations, this is called ► color constancy. Color constancy is a mystery of perception. Therefore, an important challenge in skin detection is to represent the color in a way that is invariant or at least insensitive to changes in illumination. As will be discussed shortly, the choice of the color space affects greatly the performance of any skin detector and its sensitivity to change in illumination conditions. Another challenge comes from the fact that many objects in the real world might have skin-tone colors. For example, wood, leather, skin-colored clothing, hair, sand, etc. This causes any skin detector to have many false detections in the background if the environment is not controlled.

A Framework for Skin Detection

Skin detection process has two phases: a training phase and a detection phase. Training a skin detector involves three basic steps:

1. Collecting a database of skin patches from different images. Such a database typically contains skin-colored patches from a variety of people under different illumination conditions.
2. Choosing a suitable color space.
3. Learning the parameters of a skin classifier.

Given a trained skin detector, identifying skin pixels in a given image or video frame involves:

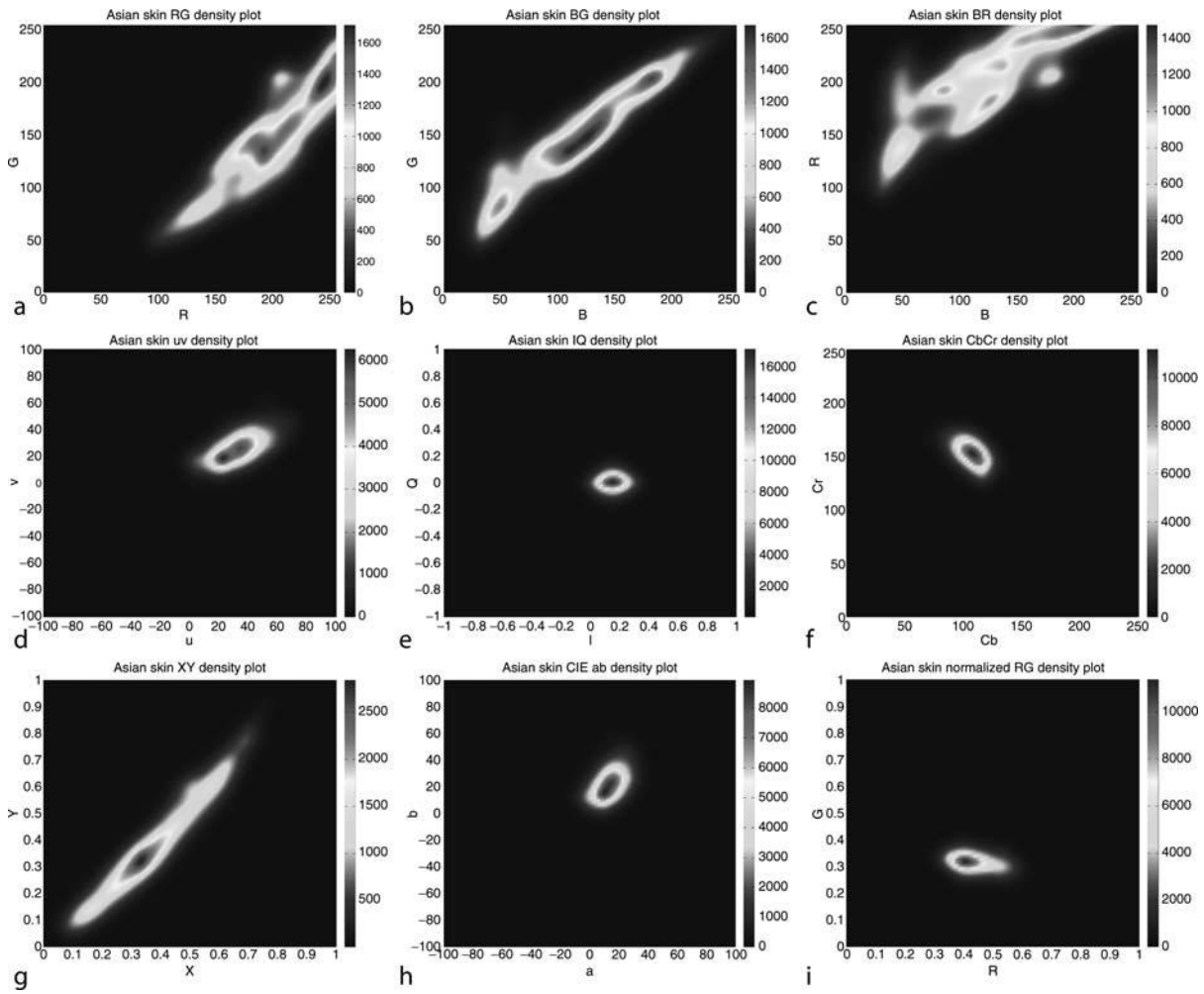
1. Converting the image into the same color space that was used in the training phase.
2. Classifying each pixel using the skin classifier as either a skin or a nonskin.
3. Typically post processing is needed using morphology to impose spatial homogeneity on the detected regions.

In any given color space, skin color occupies a part of such a space, which might be a compact or large region in the space. Such region is usually called the skin color cluster. A skin classifier is a one-class or two-class classification problem. A given pixel is classified and labeled whether it is skin or nonskin given a model of the skin color cluster in a given color space. In the context of skin classification, true positives are skin pixels that the classifier correctly labels as skin. True negatives are nonskin pixels that the classifier correctly labels as nonskin. Any classifier makes errors: it can wrongly label a nonskin pixel as skin or a skin pixel as a nonskin. The former type of errors is referred to as false positives (false detections) while the later is false negatives. A good classifier should have low false positive and false negative rates. As in any classification problem, there is a tradeoff between false positives and false negatives. The more loose the class boundary, the less the false negatives and the more the false positives. The tighter the class boundary, the more the false negatives and the less the false positives. The same applies to skin detection. This makes the choice of the color space extremely important in skin detection. The color needs to be represented in a color space where the skin class is most compact in order to be able to tightly model the skin class. The choice of the color space directly affects the kind of classifier that should be used.

Skin Detection and Color Spaces

As was highlighted by Forsyth and Fleck [6] the human skin color has a restricted range of hues and is not deeply saturated, since the appearance of skin is formed by a combination of blood (red) and melanin (brown, yellow). Therefore, the human skin color does not fall randomly in a given color space, but clustered at a small area in the color space. But it is not the same for all the color spaces. A variety of color spaces has been used in skin detection literature with the aim of finding a color space where the skin color is invariant to illumination conditions. The choice of the color spaces affects the shape of the skin class, which affects the detection process. Here, some color spaces, which are typically used in skin detection, are briefly described, and the way they affect the skin detection is discussed. The goal of the discussion is to highlight answers to the following questions: Given a skin patch, where will it be located in a given color space? Given a skin patch, what effect will changing the illumination intensity have in its location in a given color space? Given skin patches from different people from the same race, how are all these patches related in a given color space? Given skin patches from different people races, how are all these patches related in a given color space? Figures 1 and 2 help illustrate the answers for these questions. Figure 1 shows density plots for skin-colored pixels obtained from images of different Asian people plotted in different color spaces. Figure 2 shows density plots for skin-colored pixels from different people from different races: Asian, African, and Caucasian plotted in different color spaces.

RGB Color Space and Skin Detection: **RGB** color space is the most commonly used color space in digital images. It encodes colors as an additive combination of three primary colors: red (R), green (G), and blue (B). **RGB** Color space is often visualized as a 3D cube where R, G, and B are the three perpendicular axes. One main advantage of the **RGB** space is its simplicity. However, it is not perceptually uniform, which means distances in the **RGB** space do not linearly correspond to human perception. In addition, **RGB** color space does not separate ▶ luminance and ▶ chrominance, and the R,

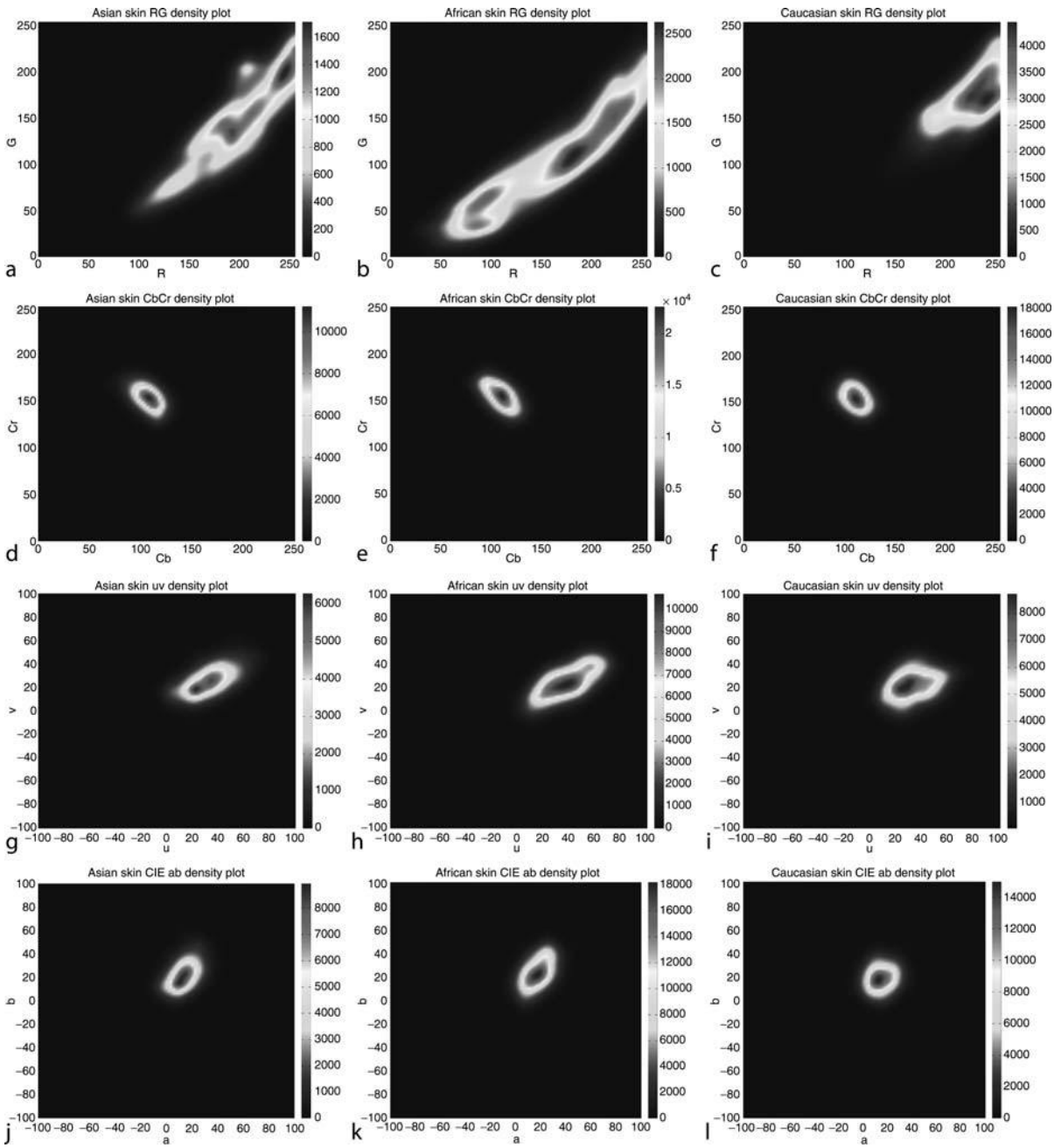


Skin Detection. **Figure 1** Density plots of Asian skin in different color spaces.

G, and B components are highly correlated. The luminance of a given **RGB** pixel is a linear combination of the R, G, and B values. Therefore, changing the luminance of a given skin patch affects all the R, G, and B components. In other words, the location of a given skin patch in the **RGB** color cube will change based on the intensity of the illumination under which such patch was imaged! This results in a very stretched skin color cluster in the **RGB** color cube. This can be noticed in the first row of Fig. 1 where skin patches from images of Asian people taken at arbitrary random illumination are plotted in the **RGB** space. The skin color cluster is extended in the space to reflect the different illumination intensities in the patches. Similarly, the skin color clusters for patches from different races will be located at different locations in the **RGB** color space. This can be seen in the first row of Fig. 2.

Despite these fundamental limitations, **RGB** is extensively used in skin detection literature because of its simplicity. For example, **RGB** is used by Rehg and Jones [8] and yields quite a satisfying performance.

TV Color Spaces and Skin Detection: A different class of color spaces is the orthogonal one used in TV tansmission. This includes **YUV**, **YIQ**, and $\mathbf{YC}_b\mathbf{C}_r$. **YIQ** is used in NTSC TV broadcasting while $\mathbf{YC}_b\mathbf{C}_r$ is used in JPEG image compression and MPEG video compression. One advantage of using these color spaces is that most video media are already encoded using these color spaces. Transforming from **RGB** into any of these spaces is a straight forward linear transformation [5]. All these color spaces separate the illumination channel (**Y**) from two orthogonal chrominance channels (**UV**, **IQ**, $\mathbf{C}_b\mathbf{C}_r$). Therefore, unlike **RGB**, the location of the skin color in the chrominance channels will not be



Skin Detection. **Figure 2** Density plots of Asian, African and Caucasian skin in different color spaces.

affected by changing the intensity of the illumination. In the chrominance channels the skin color is typically located as a compact cluster with an elliptical shape. This can be seen in Figs. 1d–f. This facilitates building skin detectors that are invariant to illumination intensity and that use simple classifiers. The density of the skin color over the chrominance channels can be easily approximated using a multivariate Gaussian distribution. Moreover, the skin colors of different races almost colocate in the chrominance channels. This can be seen in the second and third rows of Fig. 2. Therefore, using such color spaces results in skin detectors which are invariant to human race. The simplicity of the transformation and the invariant properties made such spaces widely used in skin detection applications [1, 2, 9–11, 14].

Perceptual Color Spaces and Skin Detection: Perceptual color spaces, such as **HSI**, **HSV**/**HSB**, and **HSL** (**HLS**), have also been popular in skin detection. These color spaces separate three components: the hue (H), the saturation (S), and the brightness (I,V or L). Essentially, **HSV**-type color spaces are deformations of the **RGB** color cube and they can be mapped from the **RGB** space via a nonlinear transformation. One of the advantages of these color spaces in skin detection is that they allow users to intuitively specify the boundary of the skin color class in terms of the hue and saturation. As **I**, **V** or **L** give the brightness information, they are often dropped to reduce illumination dependency of skin color. These spaces have been used by Shin et al. [11] and Albiol et al. [2].

Colorimetric Color Spaces and Skin Detection: Separating the chromaticity from the brightness is also achieved in Colorimetric color spaces, such as **CIE**-**XYZ**, **CIE**-*xy*, **CIE**-*Lab* defined by the International Commission on Illumination (Commission Internationale d'Èclairage – CIE). **CIE**-**XYZ** color space is one of the first mathematically defined color spaces (defined in 1920s). It is based on extensive measurements of human visual perception, and serves as a foundation of many other colorimetric spaces. **CIE**-**XYZ** can be achieved through a linear coordinate transformation of the **RGB** color space. The Y component corresponds to the lightness of the color (the luminance). The chromaticity values (x, y) can be achieved by central projection into the plane $X+Y+Z=1$ and then projecting into the XY plane. For details see [5]. The result is the well-known horse-shaped **CIE**-*xy* chromaticity diagram defining the hue and saturation of any color. One of the disadvantages of the **XYZ** and the *xy* color spaces is that the color differences are not perceived equally in different regions of the color space. In contrast, the **CIE**-*Lab* separates a luminance variable L from two perceptually uniform chromaticity variables (a, b). Fig. 1h shows the skin color density for Asian skin in the a,b chromaticity space. Figure 2 (last row) shows the skin color density for different races in the a,b space. Despite the many advantages of such color spaces, they are rarely used in skin detection. This is mainly because the transformation from **RGB** is more computationally expensive than other spaces. **CIE**-**XYZ** color space was used by Shin et al. [11] in comparison with other color spaces. The chrominance *xy* plane was used by Lee and Yoo [9].

Skin Classifiers

A variety of classification techniques has been used in the literature for the task of skin classification. A skin classifier is a one-class classifier that defines a decision boundary of the skin color class in a feature space. The feature space in the context of skin detection is simply the color space chosen. Any pixel which color falls inside the skin color class boundary is labeled as skin. Therefore, the choice of the skin classifier is directly induced by the shape of the skin class in the color space chosen by a skin detector. The more compact and regularly shaped the skin color class, the more simple the classifier.

The simplest way to decide whether a pixel is skin color or not is to explicitly define a boundary. Brand and Mason [4] constructed a simple one-dimensional skin classifier: a pixel is labeled as a skin if the ratio between its R and G channels is between a lower and an upper bound. They also experimented with one-dimensional threshold on IQ plane of YIQ space where the "I" value is used for thresholding. Other methods explicitly define the skin color class boundary in a two-dimensional color space using elliptical boundary models [9]. The parameters of the elliptical boundary can be estimated from the skin database at the raining phase.

Baysian Approach for Skin Detection: Skin classification can be defined probabilistically as: given a pixel with color c what is the probability of it being skin pixel $P(skin|c)$. Once this probability is computed, the pixel is labeled as a skin pixel if such probability is larger than a threshold and nonskin otherwise. Obviously such probabilities cannot be computed for every possible color (e.g., in 24 bit **RGB**, there are 256^3 colors). Fortunately, using Bayes rule, this can be rewritten as

$$P(skin|c) = \frac{P(c|skin)P(skin)}{P(c|skin)P(skin) + P(c|notskin)P(notskin)}$$

Bayes rule defines the posterior probability of a pixel being skin given its color ($P(skin|c)$) in terms of the likelihood of observing such color given the skin class ($P(c|skin)$) and the prior probability of the skin class $P(skin)$. The prior probability measures our guess about a random pixel being a skin without observing its color. The denominator in the Bayes rule is the total probability of observing the color c, a factor that does not affect the decision whether a pixel ought to be labeled as skin or nonskin. Given Bayes rule, the skin classification reduces to computing the likelihood term, i.e., $P(c|skin)$. Given a database of skin-colored pixels, the

probability density function (pdf) of $P(c|skin)$ can be estimated. Several approaches have been introduced to compute this pdf including the use of histograms [8], the use of a single Gaussian model, or a Mixture of Gaussians model [12] to approximate such pdf.

The skin classifier can also be posed as a two-class problem. From Bayes rule, this results in computing the likelihood ratio of observing a given color given a skin class versus a nonskin class, i.e., $P(c|skin)/P(c|notskin)$. Such a ratio can then be thresholded to decide whether a pixel is a skin or nonskin pixel. Besides modeling the likelihood of an observed color given the skin class, the complementary class needs to be models. That is, modeling the probability density function of nonskin pixels $P(c|notskin)$. Rehg and Jones [8] approximated such pdfs using 3D histograms in the **RGB** space based on a large database of skin and nonskin images.

Skin Detection Applications and Examples

Human face localization and detection is the first step in obtaining face biometrics. Skin color is a distinguishing feature of human faces. In a controlled background environment, skin detection can be sufficient to locate faces in images. As color processing is much faster than processing other facial features, it can be used as a preliminary process for other face detection techniques [10]. Skin detection has also been used to locate body limbs, such as hands, as a part of hand segmentation and tracking systems, e.g., [7].

Forsyth and Fleck [6] demonstrated that skin filter can be used as part of the detection process of images with naked or scantily dressed people. Their technique has three steps. First, a skin filter, based on color and texture, was used to select images with large areas of skin-colored pixels. Then, the output is fed into a geometric filter which identifies the skin-colored regions with cylindrical shapes. Those skin-colored cylinders are grouped into possible human limbs and connected groups of limbs. Images containing sufficiently large skin-colored groups of possible limbs are then reported as containing naked people.

Zheng et al. [14] presented an adaptive skin detector for detecting naked pictures on the internet. Their technique applies a face detector on the picture first to find the skin color. They argued that as skin color highly depends on illumination and the race of the person, it is more appropriate to get the skin color from the face of the person in the image. Using the skin color and the property of the texture from the detected face region, the rest of skin pixels in the image can be detected.

Skin Detection Performance

Regardless of the choice of the color space and the classification method, most published research on skin detection reports about 95% true detection while the false detection rates varies from 15 to 30%.

Summary

Skin detection in color images and videos is a very efficient way to locate skin-colored pixels, which might indicate the existence of human faces and hands. However, many objects in the real world have skin-tone colors, such as some kinds of leather, sand, wood, fur, etc., which might be mistakenly detected by a skin detector. Therefore, skin detection can be very useful in finding human faces and hands in controlled environments where the background is guaranteed not to contain skin-tone colors. Since skin detection depends on locating skin-colored pixels, its use is limited to color images, i.e., it is not useful with gray-scale, infrared, or other types of image modalities that do not contain color information. There has been extensive research on finding human faces in images and videos using other cues such as finding local facial features or finding holistic facial templates [13]. Skin detection can also be used as an efficient preprocessing filter to find potential skin regions in color images prior to applying more computationally expensive face or hand detectors.

Related Entries

- ▶ Face Localization
- ▶ Face Recognition, Overview
- ▶ Hand Recognition

References

1. Abdel-Mottaleb, M., Elgammal, A.: Face detection in complex environments from color lmages. In: Proceedings of the International Conference on Image Processing (ICIP), pp. 622–626 (1999)

2. Albiol, A., Torres, L., Delp, E.: Optimum color spaces for skin detection. In: Proceedings of the International Conference on Image Processing (ICIP), pp. I: 122–124 (2001)
3. Bernhard Fink, K.G., Matts, P.J., Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. Evol. Hum. Behav. **27**(6), 433–442 (2006)
4. Brand, J., Mason, J.: A comparative assessment of three approaches to pixellevel human skin-detection **1**, 1056–1059 (2000)
5. Burger, W., Burge, M.: Digital Image Processing, an Algorithmic Introduction Using Java. Springer, Berlin (2008)
6. Fleck, M.M., Forsyth, D.A., Bregler, C.: Finding naked people. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 593–602 (1996)
7. Imagawa, K., Lu, S., Igi, S.: Color-based hands tracking system for sign language recognition. In: FG '98: Proceedings of the Third International Conference on Face and Gesture Recognition, p. 462. IEEE Computer Society, Washington, DC, USA (1998)
8. Jones, M.J., Rehg, J.M.: Statistical color models with application to skin detection. Int. J. Comput. Vision (IJCV) **46**(1), 81–96 (2002). URL citeseer.ist.psu.edu/jones99statistical.html
9. Lee, Y., Yoo, S.I.: An elliptical boundary model for skin color detection. In: Proceedings of the International Conference on Imaging Science, Systems, and Technology (2002)
10. Senior, A., Hsu, R.L., Mottaleb, M.A., Jain, A.K.: Face detection in color images. IEEE Trans. Pattern Analy. Mach. Intell. (PAMI) **24**(5), 696–706 (2002). DOI http://dx.doi.org/10.1109/34.1000242
11. Shin, M.C., Chang, K.I., Tsap, L.V.: Does colorspace transformation make any difference on skin detection? In: WACV '02: Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, p. 275. IEEE Computer Society, Washington, DC, USA (2002)
12. Yang, M., Ahuja, N.: Gaussian mixture model for human skin color and its application in image and video databases. In: Proceedings of the SPIE: Conference on Storage and Retrieval for Image and Video Databases (SPIE 99), vol. 3656, pp. 458–466 (1999)
13. Yang, M., Kriegman, D., Ahuja, N.: Detecting faces in images: A survey. IEEE Trans. Pattern Analy. Mach. Intell. (PAMI) **24**(1), 34–58 (2002)
14. Zheng, Q.F., Zhang, M.J., Wang, W.Q.: A hybrid approach to detect adult web images. In: PCM (2), pp. 609–616 (2004)

Skin Print

▶ Skin Texture

Skin Spectroscopy

Dong Yi, Weilong Yang, Stan Z. Li
Center for Biometrics and Security Research & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China

Synonyms

Spectral analysis of skin

Definition

Skin spectroscopy is the study of interaction between radiation and human skin, as a function of wavelength. Human skin has a complicated multilayered structure and each layer is composed of different chemical substance. The determination of the interaction can be made by measuring the reflectance property of skin.

Introduction

The interaction between electromagnetic radiation and matter at various wavelengths can be used to reveal the structure of matter. The study of this interaction is named as spectroscopy. Since the composition of the object has different response under different radiation, when the surface of an object is illuminated by the radiation with different wavelength, the reflectance will vary as well.

The study of skin spectroscopy originated in the field of photobiology, which has tens of years history. From the reflectance of skin under the radiation of some wavelength, clinicians can obtain the composition of skin and observe the changes in skin or blood. In [1], Anderson et al. established an effective optical model for each layer of human skin by Kulberka-Munk approach. The work developed a quantitative, general model for the radiation transfer in the human skin. Based on this model, the optical parameters of each skin layer can be acquired.

Skin is a good basis for establishing the biometric identity, because everyone has unique skin properties in terms of color, appearance, texture, and inner structure. Skin spectroscopy provides an efficient way to

describe these properties of skin under the radiation of different wavelengths. Measurements of skin optics under radiations of multiple wavelengths can provide more information than using a single wavelength, which can enhance the performance of biometric system. In addition to the traditional visible light imaging, near-infrared and thermal-infrared imaging have been applied in the biometric system.

Spectral Bands

Skin spectroscopy is focused on the interaction between electromagnetic wave and human skin. From long to short wavelength, electromagnetic wave can be divided into several categories: radio wave, infrared (IR), visible light (VIS), ultraviolet (UV), x-ray,and gamma ray. Radio wave is usually passed around the human body with no interaction with the human skin, due to its long wavelength and low energy. X-ray and gamma ray can easily penetrate human skin to damage cells in tissue. Although ultraviolet is widely used in photo-medicine for disease diagnosis, high-intensity ultraviolet can cause skin suntan, burn, or even skin cancer. For noninvasive biometric applications, VIS and IR are two main bands that have been used practically. The spectrum of visible light is usually divided into three channels: red, green, and blue. The infrared portion of electromagnetic wave can be divided into four spectral regions: near infrared (NIR), short-wave infrared, thermal infrared (TIR), and far infrared (FIR). Fig. 1 illustrates UV, VIS, and IR bands in different wavelength.

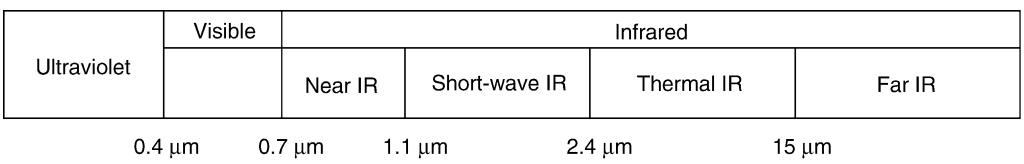
Human Skin Structure

Skin is the largest organ of the human body, which covers the whole body and protects the internal tissue from outside damage. From one's skin, we can access his mood, health condition, and attractiveness. Human skin has a complex, multilayered structure and chemical composition. Generally, from outside to inside, human skin is composed of the following layers:

- The epidermis, which is the exterior layer of skin. It does not contain blood vessels. The main type of cells in the epidermis are keratinocytes, melanocytes, Langerhans cells, and Merkels cells. Among these cells, melanocytes are the most important cells for skin spectroscopy, because they can synthesize melanin.
- The dermis, which is the layer of skin beneath the epidermis. It contains hair follicles, sweat glands, sebaceous glands, apocrine glands, lymphatic vessels, blood vessels, and nerve endings. The blood vessels provide nutrition for the epidermis and the dermis. Nerve endings can provide the sense of touch and heat.
- The subcutaneous tissue, which has no evident boundary with the dermis, can isolate body from heat and store energy. The main constituent of subcutaneous tissue is adipose.

Human Skin Optics

The radiation first interacts with the surface of the skin and then penetrates inside, but it usually cannot reach the subcutaneous tissue. Thus, the skin surface, the epidermis, and the dermis are three main aspects that been studied in skin spectroscopy. In the following sections, the optical properties of interface between air and the stratum corneum, the epidermis, and the dermis will be described. The optics of skin is dependent on the wavelength and the dose of the incident light. Due to the high energy, x-ray and gamma ray can cause chemical reactions with the cells of the skin and the body, which is called "radiolysis". Here, only the spectrum from 0.25 to 3 μm (including UV, VIS and IR) is discussed, because the other spectra do not have much significance for biometrics.



Skin Spectroscopy. **Figure 1** The electromagnetic wave radiation spectrum bands.

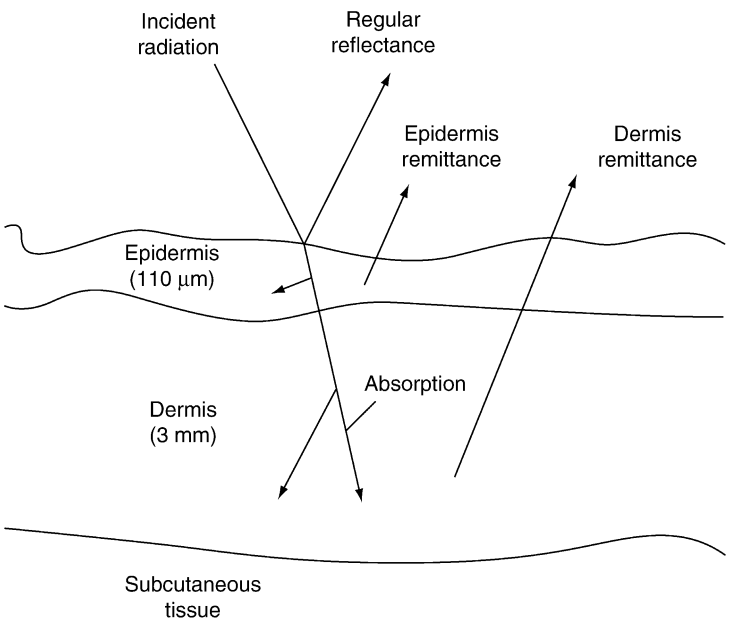
A typical process of interaction between radiation and the skin is illustrated in Fig. 2. Under the incident radiation, a small fraction of radiation is reflected by the interface between air and the stratum corneum. Generally, the reflectance is always between 4% and 7% over the entire spectrum from 0.25 to 3 μm, for all kinds of skin (white, black, etc.). Since the surface of the stratum corneum is not flat and smooth, the reflectance from skin is not *specular* except those reflected by oil. The remaining 93–96% of the incident radiation penetrates into the skin and will be absorbed or scattered.

In the UV band, there are many cells and chemical constituents in the epidermis, affecting the transmittance and the remittance of the radiation. For example, aromatic amino acids tryptophan and tyrosine are with a minimum transmittance near 0.275 μm; nucleic acids are with an absorption maximum near 0.26 μm; and numerous small aromatic molecules are with an absorption maximum at 0.277 μm, and so on. The content and distribution of melanin play an important role in determining the optical properties of epidermis.

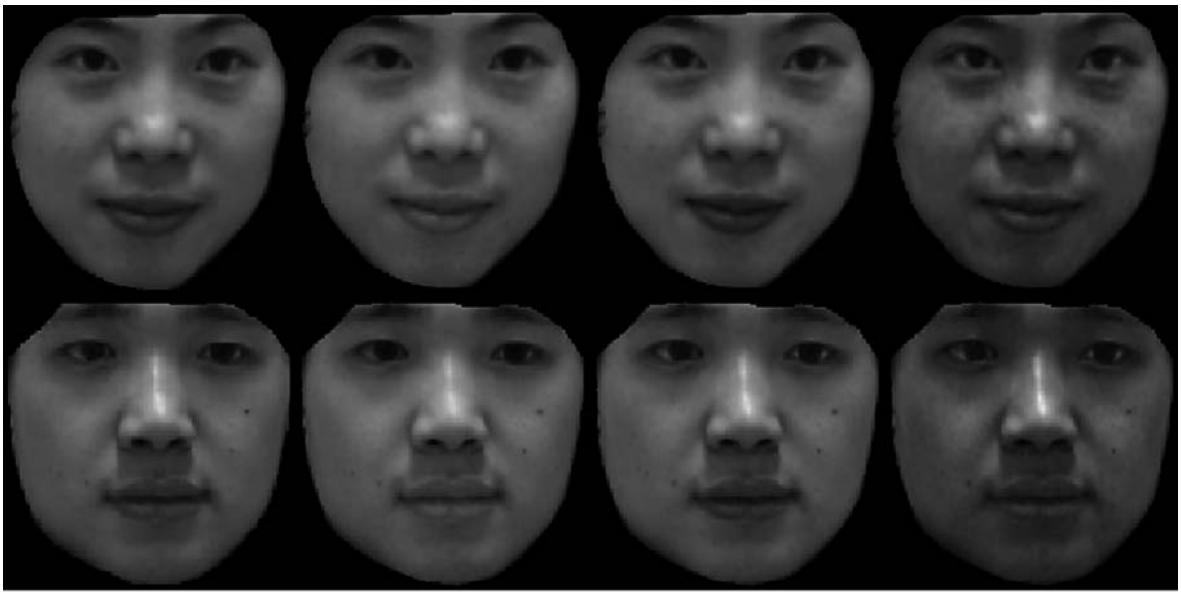
In the VIS band, melanin is the only pigment affecting the transmittance of epidermis. Shorter wavelengths, such as blue light, are highly absorbed by certain tissue components such as melanin and blood compared with longer wavelengths. In addition, the optical scattering increases as the wavelength gets shorter in this spectral range. For these reasons, longer wavelengths (red) penetrate the tissue deeper than shorter wavelengths. This phenomenon is illustrated in Fig. 3, in which the R channel is shown to be the smoothest channel and B channel contains most details.

The skin color (VIS spectrum) is affected by the type and quantity of melanin in the epidermis. This results in various skin colors from white to brown and black, although the structure of human skin is similar across different races. The absorption of melanin decreases monotonously from the short wavelength of 0.25 μm in the UV band (through the VIS band wavelength) to the long wavelength of 1.1 μm in the NIR band. Beyond 1.1μm, the absorption of melanin is negligible and both the transmittance and remittance of skin are uncorrelated with melanin. TIR of wavelength between 2.4 and 15 μm is often used passively to determine the temperature of skin.

The dermis has a significantly different structure and composition with epidermis, which causes their different optical properties. Maybe, because the dermis is deeper than epidermis, fewer studies have concentrated on dermal optics. In the dermis, scattering holds the dominant position. Many observations show that the transmittance and remittance are close to 100% across the spectrum from 0.3 to 2.4 μm, indicating that very little radiation is absorbed there. Light of longer wavelengths can penetrate the dermis deeper than shorter wavelengths.



Skin Spectroscopy. **Figure 2** The general model of interaction between radiation and skin.



Skin Spectroscopy. Figure 3 Two face images in VIS and their RGB channels.

Biometrics Applications

From the foregoing discussion, we can find out two important properties of skin spectroscopy: uniqueness and diversity, which make the skin spectroscopy become a valuable tool for the biometric technology. The uniqueness is that everyone has his unique spectroscopic properties of skin, which provides a discriminative feature for the biometric systems. The diversity is that the properties of different skin parts are diverse under the radiation of different wavelengths, which provides abundant features to describe skin.

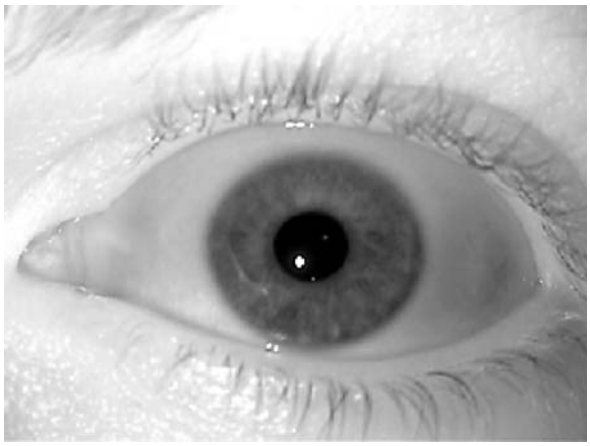
Due to the nutritious texture information, face, fingerprint, and palmprint are the three main parts of skin used in biometric systems, which are called "modality" in biometrics. In addition, their physical formation is stable and all of them can be easily imaged by the optical sensors with different wavelengths. Then, the application of skin spectroscopy in the biometric systems is introduced in the following section.

In terms of the spectrum of radiation, VIS, NIR, and TIR are the three common bands used widely in face recognition. Fig. 4 shows VIS, NIR, and TIR (thermogram) face images. Apparently, TIR face image contains less texture information than VIS and NIR images. The most conventional face recognition is based on face images captured by common VIS camera. More recently, techniques based on NIR [2, 3] and TIR [4, 5] images are introduced to overcome problems arising from changes in ambient illumination. Notably, facial thermogram is not based on skin spectroscopy, which shows the active TIR radiation of live human skin.

In fingerprint recognition, there are three common data acquiring devices: optical, ultrasonic, and capacitance. The first one belongs to spectroscopy-based methods. Optical imaging devices usually capture a digital image of fingerprint, using VIS light. One disadvantage of this type of sensor is that the imaging capacity is affected by the skin quality of the finger. Novel sensing techniques such as multispectral imaging [7, 8] have been developed to overcome this problem. The principle of the ultrasonic device is very similar to that of skin spectroscopy-based methods while just using ultrasonic instead of electromagnetic wave. Ultrasonic sensors use very high-frequency sound waves to penetrate the epidermal layer of the skin. Since the dermal skin layer exhibits the same characteristic pattern of the fingerprint, the reflected wave measurements can be used to form an image of the fingerprint. This eliminates the need for clean, undamaged epidermal skin and a clean sensing surface.

For palmprint biometric, a large region of palm supplies plenty of line patterns that can be easily captured by a low-resolution imaging device. Palmprint images are often captured by an active VIS lighting imaging device in a semiclosed environment [9].

Skin Spectroscopy. **Figure 4** VIS, NIR and TIR face images. The TIR image is from the Equinox database [6].



Skin Spectroscopy. **Figure 5** A NIR iris image.

Strictly, the iris is not the skin, but it has a rich texture like skin that can be used to distinguish different people. Its complex patterns contain many distinctive features. NIR wavelengths can penetrate the iris as deep as they penetrate the skin. Hence, NIR is used for clear and unobtrusive imaging at a distance of up to 1m further. Moreover, under active NIR light, even darkly pigmented irises reveal rich features. Fig. 5 is a NIR iris image that presents complex iris texture pattern.

Moreover, compared with these single-modal systems described earlier, hybrid biometric systems can achieve better performance by fusing some modalities or spectrum, for example, face+fingerprint [10], face+iris [11], face+palmprint [12], multispectral NIR face [2] etc.

Summary

From the perspective of skin spectroscopy, the human skin has its reflectance and absorption characteristics as a function of wavelength of illumination. Since the composition of skin across people have significant difference, skin spectroscopy has been widely used in biometrics, in which many parts of skin with nutritious texture (face, fingerprint, palmprint, and other modalities) are often chosen as objects of study. For every modality, one or many appropriate bands can be chosen according to the skin spectroscopy. By fusing some modalities or spectrum, lots of hybrid biometrics have been generated. Generally, compared with single biometric, multispectral or multimodal features of skin can lead to more effective biometric systems. More applications of skin spectroscopy to biometric technologies can be foreseen.

Related Entries

- ▶ Face Recognition
- ▶ Fingerprint Recognition
- ▶ Multibiometrics
- ▶ Multi-Spectral Biometrics
- ▶ Near Infrared
- ▶ Near Infrared Face Recognition
- ▶ Palmprint Recognition

References

1. Anderson, R.R., Parrish, J.A.: The optics of human skin. J. Invest. Dermatol. **77**, 13–19 (1981)
2. Pan, Z.H., Healey, G., Prasad, M., Tromberg, B.: Face recognition in hyperspectral images. IEEE Trans. Pattern Anal. Mach. Intell. **25**(12), 1552–1560 (2003)
3. Li, S.Z., Chu, R., Liao, S., Zhang., L.: Illumination invariant face recognition using near-infrared images. IEEE Trans. Pattern Anal. Mach. Intell. **26**(Special issu e on Biometrics: Progress and Directions) (2007)
4. Selinger, A., Socolinsky, D.A.: Face recognition in the dark. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Washington, DC, USA, pp. 129–129 (2004)
5. Chen, X., Flynn, P.J., Bowyer, K.W.: Infra-red and visible-light face recognition. Comput. Vis. Image Underst. **99**, 332–358 (2005)
6. Miezianko, R.: Terravic Research Infrared Database. http://www.terravic.com/research/index.htm (Accessed June 17, 2008)
7. Rowe, R.K., Corcoran, S.P., Nixon, K.: Biometric identity determination using skin spectroscopy. Tech. rep., http://www.lumidigm.com (2004)
8. Jain, A.K., Flynn, P., Ross, A.A. (eds.): Handbook of Biometrics. Springer (2008)
9. Zhang, D., Kong, W.K., You, J., Wong, M.: Online palmprint identification. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1041–1050 (2003)
10. Hong, L., Jain, A.: Integrating faces and fingerprints for personal identification. IEEE Trans. Pattern Anal. Mach. Intell
11. Wang, Y., Tan, T., Jain, A.K.: Combining face and iris for identity verification. In: Proceedings of International Conference on Audio- and Video-based Biometric Person Authentication. Guildford, UK (2003)
12. Gao, Y., M.Maggs: Feature-level fusion in personal identification. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA (2005)

Skin Texture

Xiangxin Zhu, Zhen Lei, Stan Z. Li
Biometrics and Security Research & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China

Synonyms

Skin Print

Definition

Generaltly, *skin texture* is the surface texture pattern of any part of human body with bare skin (e.g., face, hand, and palm). In the context of biometrics, this term commonly refers to the technologies and methods of face recognition using highly detailed facial skin texture in *high-resolution images*.

Introduction

Since the success of local features in face recognition in late 1990s, researchers have been seeking for more detailed representations of human faces. Skin texture contains plentiful detailed local information, and therefore starts to attract research attention [1]. A person's skin texture pattern is, to some extent, a unique physical trait and is distinguishable from those of others, and thus can be used for biometric identification.

The primary difference between conventional face recognition methods and skin texture methods is the resolution of facial images used. Typically, in conventional face recognition methods, the faces are scaled to 30–60 or so pixels between the centers of eyes. For skin texture methods, the inter-eye distance should be at least 90 pixels to obtain reasonable performance. With much higher ▶ image resolution, more sophisticated face recognition algorithms could possibly be proposed to yield better recognition accuracy.

However, with the increased facial image resolution, higher quality image capture devices, and more computing resources are needed for acquiring and processing the facial images, both expensive. In the 1990s, those were the main obstacles in the applications and developments of skin texture methods. Since 2000, high quality digital cameras and webcams are becoming more affordable, and the processing speed of personal computers is ten times faster than it was a decade ago with much lower price. Those have catalyzed the advance of skin texture technology.

Skin Texture Based Methods

The first work to create the skin texture face recognition algorithms was pioneered by Delean Vision around 2001 [2]. The Delean method uses a probabilistic image analysis method for skin texture matching.

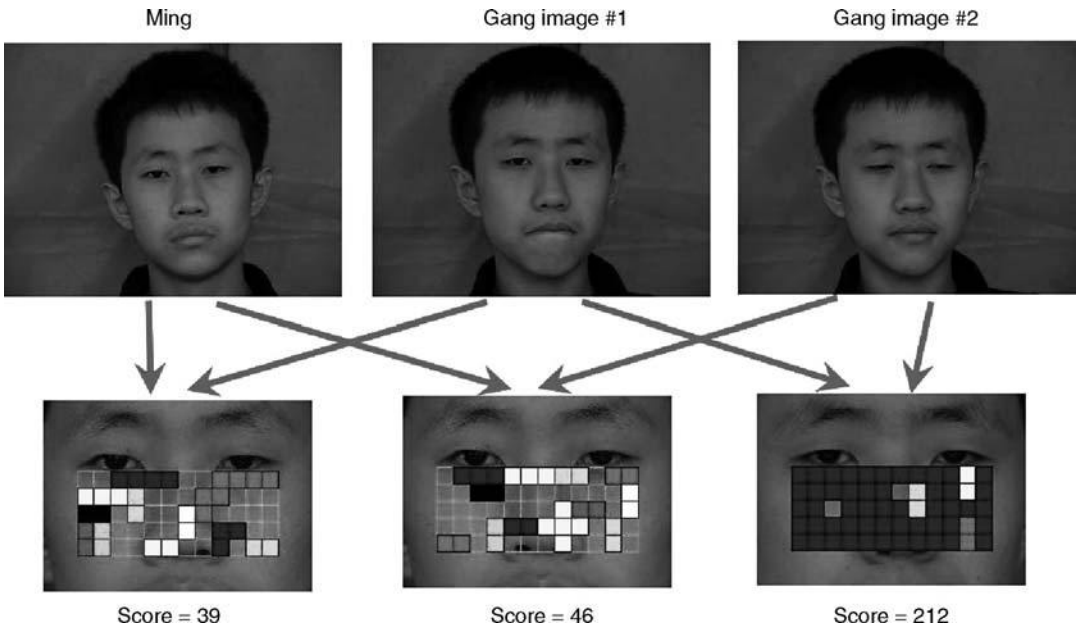
This method first defines the average luminance of each pixel as the average gray scale value of its surrounding pixels. Then the facial images are converted into binary images by comparing a pixel's gray scale value to its average luminance, with one or zero value assigned by such pixel's average luminance which is above its gray value by a given margin, as illustrated by the two feature maps in Fig. 1.

The skin area is then (usually below the eyes and above the mouth)is divided into several neighboring small blocks. For each block in the test image the algorithm searches, in a local region of its corresponding block in the reference image, for the best matching position. Figure 2 shows the matching results of identical twins, Ming and Gang, using Delean's Method. The upper row is one image of Ming and two images of Gang. The lower row gives the pairwise matching results and the corresponding similarity scores. We use different colors to mark the continuous regions after searching process. Same color indicates that the neighbhoring blocks relative position change is below a given threshold and they are considered as continuous. More continuous block pairs implies that the two images are likely to belong to a same person. The probability can be formulated as a function of the number of continuous block pairs. The underlying assumption behind this method is: With two images of a same person, the best matching positions of the neighboring blocks in the reference image tend to keep the relative position of the blocks in the test image. Otherwise, with two images of different persons, the relative position will not be kept. Finally the method takes advantage of the continuity of block pairs and the relative location of neighboring blocks to compute the similarity score (Fig. 1).

The method is usually used as a complementary to the other face matching algorithms rather than on its own. Indentix reported that the incorporation of the skin texture method into a local feature analysis (LFA) matcher [3] could increase its accuracy by 20–25%. It was also reported [4] that this method could differentiate between identical twins.

The Neven Vision method [5] uses Elastic Bunch Graph [6] to represent faces and is capable of placing this graph with high precision on a face in a presented image. Based on the found facial landmarks, corresponding to facial landmarks like eye corner or the tip of the nose, it first accurately locates the areas in the face that are used for skin texture analysis. The skin areas are then warped and normalized before matching. In the matching step, for each selected skin patch, the feature vector is matched individually to the face region of reference image and the most similar skin patch is identified. The result is a more or less distorted version of the graph in the original image. From the similarity and distortion between the original

Skin Texture. **Figure 1** Left: the blocks in a gallery skin texture image. Right: the best matching positions in a probe skin texture image.



Skin Texture. **Figure 2** An example of Delean's method in matching of twins.

graph and the matched graph, a similarity score is computed incorporating local geometric constraints [5]. Having achieved an impressive performance in FRVT2006 [7], Neven Vision also claims that their skin texture analysis is about 1,000 times faster than the Delean/Identix algorithm. This method works on very high resolution images with inter-pupil distance of at least about 600 pixels. When this condition is fulfilled, it can achieve better recognition accuracy than iris-based method and can even outperform human performance [7].

Pros and Cons

Skin texture based methods have the following advantages:

- The facial skin has a fine texture that is determined randomly during embryonic gestation. Even identical monozygotic twins have completely independent skin textures.
- The acquisition of skin texture is similar to capturing a photograph, and can be performed from several meters away. There is no need for the person to be identified to touch any equipment that has recently been touched by a stranger, thereby eliminating an objection that has been raised in some cultures against finger-print scanners, where a finger has to touch a surface, or retinal scanning, where the eye can be brought very close to a lens.
- Skin texture contains plentiful local detailed information and can achieve a very high recognition accuracy.

However, they have the following disadvantages:

- High resolution is of vital importance for skin texture based methods. The performance will dramatically drop with blurred images because the image resolution is reduced.
- Skin texture may not be the same with aging, exaggerative expression, pose variation, and illumination changes.
- In surveillance environment, it is difficult to captain high resolution face images at a far distance if the person to be identified is not cooperative by holding the head still and looking at the camera. This may narrow possible applications of skin texture methods.

Summary

As a new method for biometrics, skin texture has demonstrated its effectiveness in identifying individuals in high resolution images. The increased demand for accurate, reliable, fast, and convenient biometric

technologies, and availability of inexpensive computing resources and cameras will all benefit to the rapid technical advances in this direction. Although there are several unsolved problems and disadvantages as aforementioned, the skin texture method is surely of significance from both application and research perspectives.

Related Entries

- ► Face Recognition
- ► Local Image Features
- ► Skin Spectroscopy

References

1. Cula, O., Dana, K., Murphy, F., Rao, B.: Skin Texture Modeling. Int. J. Comput. Vis. **62**(1–2), 97–119 (2005)
2. Delean, B., Escolls, R.: Method and apparatus for probabilistic image analysis. u.s. patent application publication no. 2004/0052418 (2004)
3. Penev, P., Atick, J.: Local feature analysis: a general statistical theory for object representation. Network: Comput. Neural Syst. **7**(3), 477–500 (1996)
4. Identix Turbo charges its Biometric Engineers, Biometric Technology Today, ISSN 0969-4765/04, **12**(4), 1,12 (April 2004).
5. Adam, H., Neven, H., Steffens, J.: Single image based multi-biometric system and method u.s. patent application publication no. 2006/0050933 (2006)
6. Wiskott, L., Fellous, J., Kruger, N., Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 775–779 (1997)
7. Phillips, P., Scruggs, W., O'Toole, A., Flynn, P., Bowyer, K., Schott, C., Sharpe, M.: Frvt 2006 and ice 2006 large-scale results (2007)

Skull, Forensic Evidence of

Mingquan Zhou
Beijing Normal University, Beijing, China

Synonyms

Craniofacial reconstruction; Skull-photo superimposition; Craniofacial superimposition

Definition

Forensic evidence of skull is the technique of obtaining the identity of a person based on attributes associated with an individual skull. Generally, the methods of determining the identity of a given skull can be classified into two categories: craniofacial reconstruction, which reconstructs the facial appearance from a given skull based on the assumed relationship between the soft tissue envelope and the underlying skull substrate [1, 2] and skull-photo superimposition or craniofacial superimposition, which superimposes the skull and face photo in a certain way to verify whether they are from the same person, by comparing their feature points [3].

Introduction

In forensic pathology, it is essential to determine the identity of the dead. However, in many instances, due to the storage environment and discovery time of the corpse, the soft tissue becomes highly decomposed, and forensics cannot determine the appearance of the dead directly. As a result, attempts have been made to reconstruct the appearance from the remaining skeletons, especially the skull. Research has revealed that there are certain relationships between the facial soft tissue and the skull, and the skull has evident influence on the shape, location and structure of the facial soft tissue. Thus, the facial appearance is related to the skull to a certain extent, and it ensures the fidelity of the forensic evidence of skull [4, 5].

To determine the identity of a skull, an option is to reconstruct the 3D facial appearance explicitly or implicitly. The former is known as craniofacial reconstruction, which reconstructs the 3D facial appearance based on the given skull. Face is the most expressive part of the human being. Reconstructing a realistic 3D facial appearance is a challenging task. The latter method referred to as skull-photo superimposition does not reconstruct the 3D facial appearance explicitly. Instead, it superimposes the skull and photo at the same imaging condition to check whether the given skull and photo are of the same person, by matching their feature points. As in the case of fingerprint and tooth-based inspection, skull-photo superimposition-based inspection can also verify the identity, and is an important method in ► forensic anthropology.

In cases where the body is dismembered, skull-photo superimposition becomes even more important and unique in identity verifications. DNA can be of value in these cases and even when only the bones remain.

Craniofacial Reconstruction

Traditionally, craniofacial reconstruction methods can be classified into 2D reconstruction and 3D reconstruction. In 2D reconstruction, the hand drawn portrait of the dead is viewed by experienced anthropologists, artists or forensics based on the skull. 3D reconstruction methods normally set some marker points with height on the skull model, and use clay as a replacement of soft tissue, to shape the facial appearance based on the skull according to anatomical knowledge. However, these traditional reconstruction methods have several limitations:

1. The whole process is very time-consuming, it takes almost a month to reconstruct a person's appearance.
2. The result relies heavily on the expertise and experience of the reconstructors, the accuracy and reliability of the reconstruction is affected by many human factors.
3. The soft tissue and shape of the human face is different for each human race and living environment, it is favorable to generate several reconstructed facial appearances simultaneously when the race of a given skull is not known. Using traditional reconstruction methods, however, only one appearance for one race can be obtained.
4. It is difficult to edit or modify the reconstructed facial appearance once it is complete.

In the past decades, however, imaging technologies like ▶ CT and MRI have been applied to convert the skull into digital data so that it can be stored and processed easily, leading to computer-aided craniofacial reconstruction. In a typical computer-aided craniofacial reconstruction system, the relationship between the skull and soft tissue is not determined by the expertise and experience of the reconstructors. Instead, the relationship is learnt from the training data, which are reference pairs of captured skull data and soft tissue data. The knowledge of obtaining facial detail from the skull and the distribution law of the soft issues are obtained through feature point pairs on the skull and face appearance. This guides the creation of an individual skull. Traditionally, several methods are available to obtain the training samples of reference skull data and soft tissue data, for example, using ▶ acupuncture on the dead body to measure the thickness of the soft tissue. Following advances in imaging technologies, CT technologies may be used to collect training data for craniofacial reconstruction more accurately and efficiently. A pre-processing step, typically consisting of filtering and geometric transformation, is performed to eliminate the noise in each 2D CT slide and align the slides correctly [6, 7]. Sometimes, to facilitate 3D reconstruction of skull and face, edge detection and edge tracking is performed on the 2D CT slides to extract the contour curves in each slide [8–10]. After pre-processing, 3D reconstruction of the skull and face can be undertaken to generate the training samples for the craniofacial reconstruction system. In some systems, an explicit model describing the relationship between the skull and soft tissue is built from the training samples. Along with the development of statistical learning theory, some systems instead reconstruct 3D facial appearance directly from a given skull by using statistical learning methods, without developing an explicit model describing the knowledge of these training samples. However, though craniofacial reconstruction can be used as an auxiliary method to aid detectives in their work, its result cannot be used as legal evidence in a court of law.

Craniofacial reconstruction can also be used in ▶ palaeoanthropology, for example, to reconstruct the facial appearance of famous ancient people according to their skull, in order to see how these people looked and compare it with their portraits. Further, Craniofacial reconstruction can be used in facial surgery to help simulate the process and effect of the surgery, providing a more detailed and accurate surgery plan, and reducing the risk of operation.

Skull-photo Superimposition

Skull-photo superimposition was first developed to verify the fidelity of portraits of ancient notables. In the 1980s doubts were raised about the authenticity of many of the portraits as it was believed that the painters wanted to please these notables. To verify this conjecture, anatomists compared the skulls and

portraits of notables. However, the earliest attempts to use the skull for identity verification in literature were made at the end of the 19th century by Schau Ffhausen, Von Froliep and other scientists. In 1935 Brash recorded the first successive use of craniofacial superimposition as forensic evidence in the Ruxton case. Since then, along with the development of photographic and video technologies, many improvements and upgrades have been made by researchers. In the last 20 years, researchers have extended craniofacial superimposition from photographic superimposition to video superimposition, which has proven to be a successful craniofacial identification.

A typical skull-photo superimposition operation can be separated into the following stages: First, the skull data is obtained through a 3D scanner or reconstructed from CT scans. After collecting the 3D skull data, the difference between locations, sizes and orientations of the skull and photo need to be minimized for the superimposition task. As the photo is difficult to manipulate in three-dimension, the skull is normally edited or adjusted in three-dimension to be aligned with the photo. Moreover, the facial photo may need to exposed to several pre-processing steps such image enhancement, scaling, and rotation [6, 7]. Next, facial features including head contour curves are extracted from the facial photo [8–10]. In the next stage marker points are selected, the traditional method of manual selection relies on the expertise of operators, and the result tends to be erroneous. On the other hand, despite the recent advances in computer vision, a fully automatic method of marker points selection is not realistic at the current stage. In practice, a hybrid system of combining manually labeling and computer verification yields a reasonable result. Finally, skull-photo registration techniques and standardized verification process are performed to verify the identity [11–14]. Unlike craniofacial reconstruction, skull-photo superimposition operations, if performed well, can be used as legal evidence in court.

Related Entries

- ▶ Biometrics, Overview
- ▶ Forensic DNA Evidence
- ▶ Identification
- ▶ Verification

References

1. Prag, J., Neave, R.: Making Faces Using Forensic and Archaeological Evidence. British Museum Press, UK (1997)
2. Taylor, K.T.: Forensic Art and Illustration. CRC Press, New York (2001)
3. Dorion, R.: Photographic superimposition. JFSCA. **28**(3), 724–734 (1983)
4. Wilkinson, C.: Forensic Facial Reconstruction. Cambridge University Press, Cambridge (2004)
5. Stephan, C., Henneberg, M.: Building faces from dry skulls: Are they recognized above chance rates? J. Forensic. Sci. **46**(3), 432–440 (2001)
6. Hummel, A.D.: Image enhancement by histogram transformation. Comput. Graph Image Process. **6**(2), 184–195 (1977)
7. Ketcham, D.J.: Real-time image enhancement technique. In: Proceeding SPIE/OSA, Conference on Image Processing, pp. 120–125. Pacific Grove, CA (1976)
8. Kass, M., Witbin, A., Tetzopoulos, D.: Snake: Active contour model. Int. J. Comput. Vis. **1**(4), 321–331 (1988)
9. Lai, K.F., Chin, R.T.: Deformable contours modeling and extraction. IEEE Trans. Pattern Anal. Mach. Intell. **17**(11), 1084–1090 (1995)
10. Gunn, S.R., Nixon, M.S.: Snake head boundary extraction using global and local energy minimization. Proceeding of international conference on pattern recognition (ICPR '96), pp. 581–585. Vienna, Austria (1996)
11. Besl, P., McKay, N.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)
12. Pelizarri, C.A., Chen, G., Spelbring, D.: Accurate three-dimensional registration of CT, PET, and/or MRI images of the brain. J. Comp. Assist. Tomogr. **13**(1), 20–26 (1989)
13. Wells, W.M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. Med. Image Anal. **1**(1), 35–51 (1996)
14. Vandermeulen, D., Collignon, A., Michiels, J., Bosmans, H., Suetens, P., Marchal, G.: Multi-modality image registration within covira. Med. Image-Anal. **19**, 29–42 (1995)

Skull-Photo Superimposition

▶ Skull, Forensic Evidence of

Slap Or Four-Four-Two device

It refers to a device used to capture the ten fingerprints of a person using the following capture sequence:

simultaneous capture of the four fingers (index, middle, ring, and little finger) of one hand, simultaneous capture of the four fingers of the other hand, and simultaneous capture of the two thumbs. A segmentation algorithm is needed to detect and separate the finger in ten single images.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Smart Cameras

Smart cameras are cameras that along with the sensor have additional processing elements (micro-controllers, DSP) for processing data on-site. The results of the processing are then transmitted to other nodes in the network. Smart cameras form an important part of the distributed computing idea, as only the relevant information from each camera is transmitted to other cameras/central processing agency. This alleviates the burden on other nodes to replicate the processing as well as the need for superior computing resources at a central node.

► Surveillance

Smart Card

A smart card is a small plastic card with an embedded microchip that can store and/or process information. It can receive and submit data to or from any system equipped with an appropriate card-reader module. In commercial hand-geometry devices, for example, smart cards are often used as storage media for user-templates and as such eliminate the need for storing templates in the internal memory of the device. Its synonyms include integrated circuit card (ICC) and chip card.

► Hand-Geometry Device
► Tamper-proof Operating System

Soft Biometrics

Karthik Nandakumar[1], Anil K. Jain[2]
[1]Institute for Infocomm Research, A*STAR, Fusionopolis, Singapore
[2]Michigan State University, East Lansing, MI, USA

Synonyms

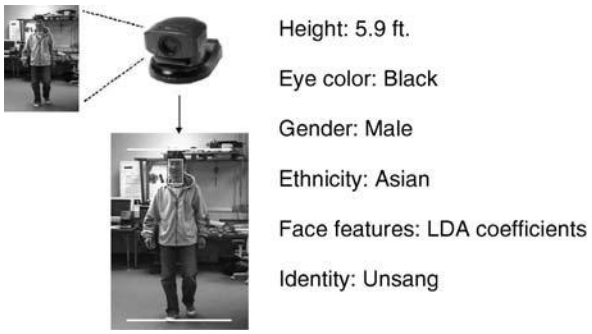
"Light" biometrics

Definition

Any anatomical or behavioral characteristic that provides some information about the identity of a person, but does not provide sufficient evidence to precisely determine the identity can be referred to as a soft biometric trait. Personal attributes like gender, ethnicity, age, height, weight, eye color, scars, marks, tatoos, and voice accent are examples of soft biometric traits. Soft biometric information complements the identity information provided by traditional (primary) biometric identifiers such as fingerprint, face, iris, and voice. Hence, utilizing soft biometric traits can improve the recognition accuracy of primary biometric systems.

Introduction

Systems that consolidate evidence from multiple sources of biometric information (e.g., face, fingerprint, hand geometry, iris, etc.) in order to reliably determine the identity of an individual are known as multibiometric systems [1]. Multibiometric systems can alleviate many of the limitations of unibiometric systems such as nonuniversality and lack of distinctiveness, thereby reducing the error rates significantly. However, using multiple biometric traits will increase the enrollment and verification times, cause more inconvenience to the users, and increase the overall cost of the system. An alternate way for reducing the error rates of the biometric system without causing any additional inconvenience to the user is to incorporate soft identifiers of human identity like gender, ethnicity, height, eye color, etc. into a (primary) biometric recognition system [2]. Figure 1 depicts a scenario where both primary (face) and soft (gender, ethnicity, height, and



Soft Biometrics. Figure 1 A scenario where the primary biometric identifier (face) and the soft biometric attributes (gender, ethnicity, eye color and height) are automatically extracted and utilized to verify a person's identity.

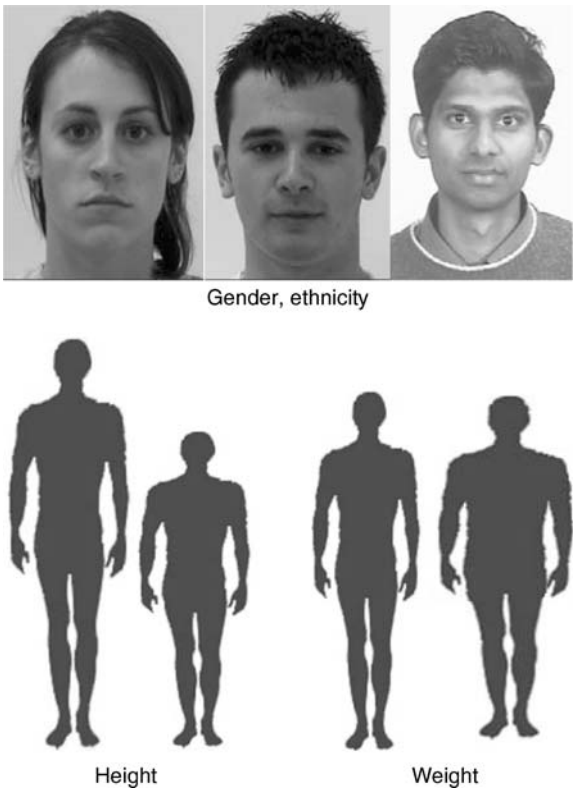
eye color) biometric information can be automatically extracted and utilized to verify an user's identity. In this scenario, the height of the user can be estimated as he approaches the camera and his age, gender, ethnicity, and eye color can be estimated from his face image. These additional attributes can be used along with the face biometric to accurately identify the person.

Motivation for Utilizing Soft Biometric Traits

The usefulness of soft biometric traits in improving the performance of the primary biometric system can be illustrated by the following example. Consider three users A (1.8 m tall, male), B (1.7 m tall, female), and C (1.6 m tall, male) who are enrolled in a fingerprint biometric system that works in the identification mode. Suppose user A presents his fingerprint sample X to the system, it is compared to the templates of all the three users stored in the database and the posteriori matching probabilities of all the three users given the sample X are calculated. Assume that the outputs of the fingerprint matcher are $P(A|X) = 0.42$, $P(B|X) = 0.43$, and $P(C|X) = 0.15$. In this case, user A will be falsely identified as user B based on the Bayesian decision rule. However, assume that as the user approaches the fingerprint sensor, there exists a secondary system that automatically identifies the gender of the user as male and measures the user's height as 1.78 m. This information, in addition to the posteriori matching probabilities given by the fingerprint matcher, is likely to lead to a correct identification of the user as user A.

The first biometric system developed by Alphonse Bertillon in 1883 used anthropometric features such as the length and breadth of the head and the ear, length of the middle finger and foot, height, etc. along with attributes like eye color, scars, and tatoo marks for ascertaining a person's identity. These measurements were obtained manually by Bertillon. Although each individual measurement in the Bertillonage system may exhibit some (intra-user) variability, a combination of several quantized (or binned) measurements was sufficient to manually identify a person with reasonable accuracy. The Bertillon system was dropped in favor of the Henry's system of fingerprint identification over 100 years back due to three main reasons: (1) lack of persistence – the anthropometric features (e.g., height) can vary significantly for juveniles; (2) lack of distinctiveness – features such as skin color or eye color cannot be used for distinguishing between individuals coming from a similar ethnic background; and (3) the huge time, effort, and training required to get reliable measurements.

While the anthropometric features used in the Bertillon system provide some information about the identity of the user, they are not sufficient for accurately identifying the user. Hence, these attributes can be referred to as "soft biometric traits". Figure 2 shows some examples of soft biometric traits. Since the soft biometric information complements the identity information provided by traditional (primary) biometric identifiers such as fingerprint, face, iris, and voice, utilizing soft biometric traits can improve the recognition accuracy of primary biometric systems. Many practical biometric systems collect soft biometric information about the users during enrollment. For example, the fingerprint card used by the Federal Bureau of Investigation (FBI) includes information on the gender, ethnicity, height, weight, eye color, and hair color of the person along with the prints of all ten fingers. However, the searches in the FBI Automated Fingerprint Identification System (AFIS) are solely based on fingerprints and the potentially useful soft biometric information is either ignored during the search or used only for manual verification after a short list of potential fingerprint matches is identified. If the soft biometric characteristics can be automatically extracted and/or utilized during the automated matching process, the overall accuracy of the system will improve and the need for manual intervention will be reduced.



Soft Biometrics. Figure 2 Examples of soft biometric traits.

Challenges in Using Soft Biometrics

Two key challenges need to be addressed to incorporate the soft biometric information into the traditional biometric framework. The first challenge is the *automatic and reliable extraction of soft biometric information* in a nonintrusive manner, without causing any inconvenience to the users. It must be noted that the failure of Bertillon-like systems was caused by the unreliability and inconvenience in the manual extraction of these features. Once the soft biometric information about a user is available, the challenge is to *optimally combine this information with the* ▶ *primary biometric identifier* so that the overall recognition accuracy is enhanced. While soft biometric traits can be used for ▶ filtering a large database or for tuning the parameters of a biometric system, such applications require a highly accurate soft biometric feature extraction module. Since it is very difficult to extract soft biometric features with 100% accuracy, the information fusion system needs to be designed in such a way that the overall recognition accuracy is enhanced even when the soft biometric feature extraction is not perfect.

Automatic Soft Biometric Feature Extraction

Soft biometric traits are available and can be extracted in a number of biometric applications. For example, attributes like gender, ethnicity, age, and eye color can be extracted with sufficient reliability from the face images. Automatic recognition of gender has been extensively studied and a majority of the gender recognition systems proposed in the literature are based on frontal face images. Furthermore, most of the these systems follow an appearance-based approach to gender recognition (see [3, 4] and the references therein). The face images are typically cropped to include only the forehead, eyes, nose, and mouth regions, and normalized for pose and illumination changes. A pattern classifier is directly trained using the normalized face images to learn the decision boundary between the male and female classes. The accuracy of face-based gender recognition systems is typically around 90% when presented with good quality frontal face images. Some of the techniques used for gender recognition can also be applied to classify people based on their ethnicity [3, 5].

Automatic age determination is a more difficult problem than gender classification. Kwon and Lobo [6] presented an algorithm for age classification from facial images based on cranio-facial changes in feature-position ratios and skin wrinkle analysis. More recently, Lanitis et al. [7] performed a quantitative evaluation of the performance of various classifiers developed for the task of automatic age estimation from face images. All the classifiers used eigenfaces obtained using ▶ principal component analysis (PCA) as the input features. Quadratic models, shortest distance classifier, neural network classifiers, and hierarchical classifier were used for estimating the age. The best age estimation algorithm had an average absolute error of 3.82 years, which was comparable to the error made by humans (3.64 years) in performing the same task. Geng et al. [8] proposed an iterative learning algorithm known as AGES for age estimation from PCA features. The AGES algorithm achieved a mean absolute error of 6.77 years in estimating the age on a database with 1,002 face images obtained from 82 subjects.

Gender, speech accent, and perceptual age of the speaker can also be inferred from the speech signal. The weight of an user can be measured by asking him to stand on a weight sensor while he is providing his primary biometric. Ailisto et al. [9] used body fat measurement as a soft biometric trait. The height of a person can be estimated from a real-time sequence of images as the user approaches the biometric system. Jain et al. [2] implemented a real-time vision system for automatic extraction of gender, ethnicity, height, and eye color. The system was designed to extract the soft biometric attributes as the person approaches the primary biometric system to present his primary biometric identifier (face and fingerprint). Their soft biometric system is equipped with two pan/tilt/zoom cameras. Camera 1 monitors the scene for any human presence based on the motion segmentation image. Once camera 1 detects an approaching person, it measures the height of the person and then guides camera 2 to focus on the person's face.

Fusion of Primary and Soft Biometric Information

Jain et al. [2] developed a Bayesian framework for fusion of primary and soft biometric features. The main advantage of this framework is that it does not require the soft biometric feature extractors to be perfect (100% accurate). Assume that the primary biometric system is based on R_p ($R_p \geq 1$) biometric identifiers like fingerprint, face, iris, and hand geometry. Further, the soft biometric system is based on R_s ($R_s \geq 1$) attributes like age, gender, ethnicity, eye color, and height. Let $\omega_1, \omega_2, \ldots, \omega_M$ represent the M users enrolled in the database. Let $x = [x_1, x_2, \ldots, x_{R_p}]$ be the collection of primary biometric feature vectors. Let $p(x_j|\omega_k)$ be the likelihood of observing the primary biometric feature vector x_j given the user is ω_k. If the output of each individual modality in the primary biometric system is a set of match scores, $s_k = [s_{1,k}, s_{2,k}, \ldots, s_{R_p,k}]$, one can approximate $p(x_j|\omega_k)$ by $p(s_j|\omega_k)$, provided the genuine match score distribution of each modality is known.

Let $y = [y_1, y_2, \ldots, y_{R_s}]$ be the soft biometric feature vector, where, for example, y_1 could be the gender, y_2 could be the eye color, etc. The posteriori probability of user ω_k given both x and y can be calculated by applying the ▶ Bayes rule as follows:

$$P(\omega_k|x, y) = \frac{p(x, y|\omega_k)P(\omega_k)}{p(x, y)}, \quad (1)$$

where $P(\omega_k)$ is the prior probability of observing user ω_k. If all the users are equally likely to access the system and if all the primary biometric feature vectors $x_1, \ldots, x_{R_p}$ and all the soft biometric variables $y_1, y_2, \ldots, y_{R_s}$ are independent of each other given the user's identity ω_k, the discriminant function, $g_k(x, y)$, for user ω_k, can be computed as

$$g_k(x, y) = \sum_{j=1}^{R_p} \log p(x_j|\omega_k) + \sum_{r=1}^{R_s} \log p(y_r|\omega_k). \quad (2)$$

During the identification phase, the input biometric sample is compared with the templates of all the M users enrolled in the database and the discriminant functions $g_1, \ldots, g_M$ are computed. The test user is identified as that user with the largest value of discriminant function among all the enrolled users. A simple method for computing the soft biometric likelihoods $p(y_r|\omega_k), r = 1, 2, \ldots, R_s, k = 1, 2, \ldots, M$ is to estimate them based on the accuracy of the soft biometric feature extractors. Jain et al. [2] also suggested the use of a scaling factor β_r, $0 \leq \beta_r \leq 1$, to flatten the likelihood distribution of each soft biometric trait. The scaling factor β_r can act as a measure of the reliability of the rth soft biometric feature and its value can be set depending on the environment (hostile or friendly) in which the system operates.

Summary

In addition to the match scores provided by the biometric matchers, ancillary information may also be available to a biometric system. Soft biometric characteristics like gender, ethnicity, height, and weight provide some information about the identity of the user. Although the soft biometric information alone is not sufficient for accurate recognition, they can be used to complement the information provided by the primary biometric identifiers like fingerprint, iris, and face. Techniques for automatically extracting soft biometric information have been developed only recently. Hence, fusion schemes that incorporate such ancillary information have not been thoroughly explored and there is a large scope for conducting more in-depth research in this area.

Related Entries

- ▶ Biometrics, Overview
- ▶ Biometric Fusion
- ▶ Face Recognition
- ▶ Fingerprint Recognition
- ▶ Multibiometrics

References

1. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. Springer, Berlin (2006)
2. Jain, A.K., Nandakumar, K., Lu, X., Park, U.: Integrating faces, fingerprints and soft biometric traits for user recognition. In: Proceedings of ECCV International Workshop on Biometric Authentication (BioAW), vol. LNCS 3087, pp. 259–269. Springer, Prague, Czech Republic (2004)
3. Gutta, S., Huang, J.R.J., Phillips, P.J., Wechsler, H.: Mixture of experts for classification of gender, ethnic origin, and pose of human faces. IEEE Trans. Neural Netw. **11**(4), 948–960 (2000)
4. Moghaddam, B., Yang, M.H.: Learning gender with support faces. IEEE Trans. Pattern Anal. Mach. Intell. **24**(5), 707–711 (2002)
5. Lu, X., Jain, A.K.: Ethnicity identification from face images. In: Proceedings of SPIE Conference on Biometric Technology for Human Identification, vol. 5404, pp. 114–123. Orlando, USA (2004)
6. Kwon, Y.H., Lobo, N.V.: Age classification from facial images. In: Proceedings of IEEE Conference on Computer Vision and Patern Recognition, pp. 762–767 (1994)
7. Lanitis, A., Draganova, C., Christodoulou, C.: Comparing different classifiers for automatic age estimation. IEEE Trans. Syst. Man Cybern. Part B: Cybern. **34**(1), 621–628 (2004)
8. Geng, X., Zhou, Z.H., Zhang, Y., Li, G., Dai, H.: Learning from facial aging patterns for automatic age estimation. In: Proceedings of the 14th Annual ACM International Conference on Multimedia, pp. 307–316. Santa Barbara, USA (2006)
9. Ailisto, H., Vildjiounaite, E., Lindholm, M., Mkel, S.M., Peltola, J.: Soft biometrics - combining body weight and fat measurements with fingerprint biometrics. Pattern Recognit. Lett. **27**(5), 325–334 (2006)

Soleprint Device

▶ Fingerprint, Palmprint, Handprint and Soleprint Sensor

Soleprint Sensor

▶ Fingerprint, Palmprint, Handprint and Soleprint Sensor

Sound

Sound is a pressure wave which is created by a vibrating object.

▶ Speech Production

Sound Generation

▶ Speech Production
▶ Voice Sample Synthesis

Sources of Evidence

▶ Sources of Information in Biometric Fusion

Sources of Information in Biometric Fusion

Arun Ross
Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV, USA

Synonyms

Sources of evidence

Definition

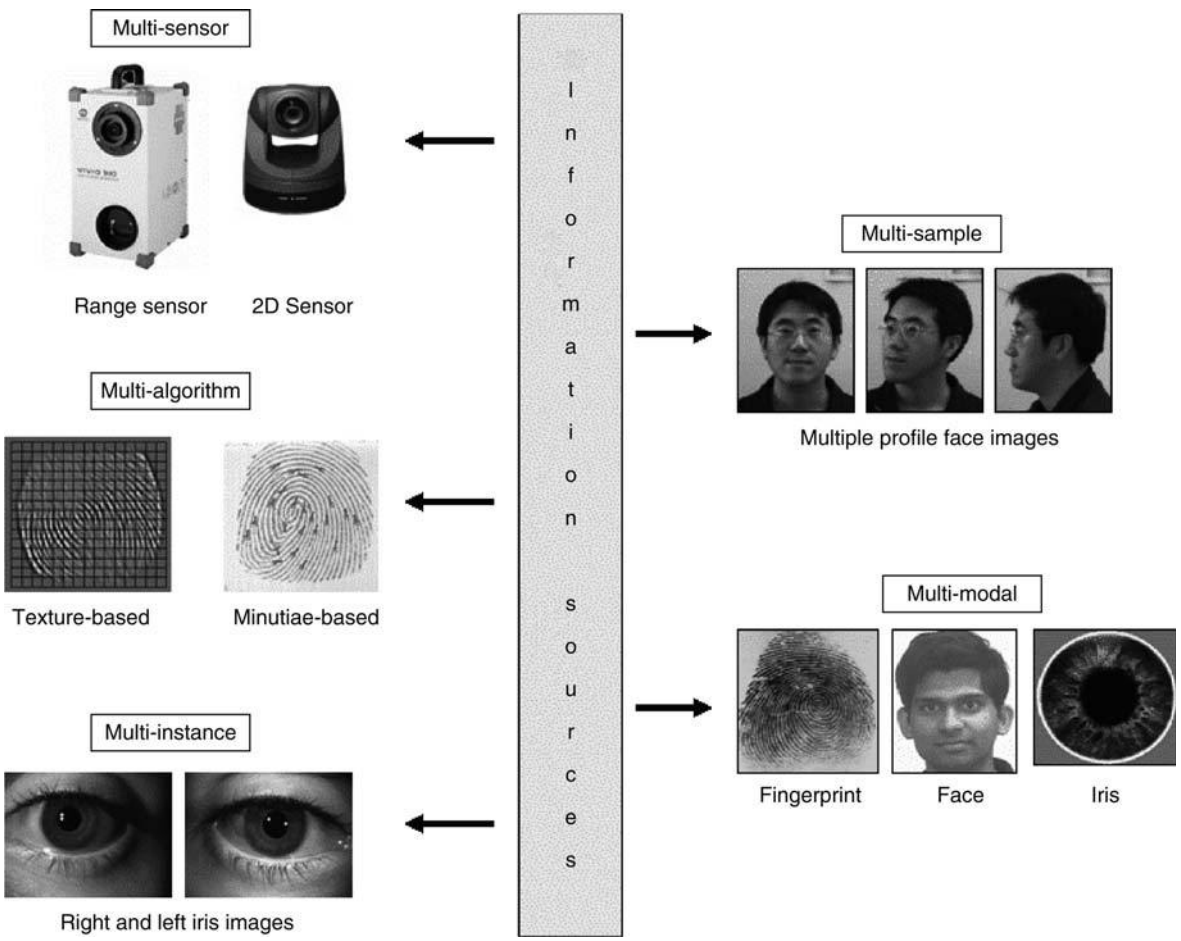
Multibiometric systems rely on the evidence presented by multiple sources of biometric information. Based on

the nature of these sources, a multibiometric system can be classified into one of the following six categories: multi-sensor, multi-algorithm, multi-instance (or multi-unit), multi-sample, multimodal, or a hybrid system. This terminology facilitates the characterization of multibiometric systems in a systematic manner. By identifying the various sources of biometric information and by understanding the challenges associated in consolidating them, appropriate fusion strategies can be devised for performing biometric fusion.

Introduction

Information fusion refers to the reconciliation of evidence presented by multiple sources of information in order to generate a decision. In the context of biometrics, evidence reconciliation plays a pivotal role in enhancing the recognition accuracy of human authentication systems and is referred to as multibiometrics. Multibiometric systems combine the information presented by multiple biometric sensors, algorithms, samples, units, or traits. Besides enhancing matching performance, these systems are expected to improve population coverage, deter spoofing, and impart fault-tolerance to biometric applications.

In order to characterize a multibiometric system, it is essential to know the various sources of information that are being consolidated. Based on the nature of these sources, a multibiometric system can be classified into one of the following six categories [1]: multi-sensor, multi-algorithm, multi-instance, multi-sample, multimodal, and hybrid (see Fig. 1). Each of these systems is described in the narrative below.



Sources of Information in Biometric Fusion. **Figure 1** Sources of information for biometric fusion.

Multi-Sensor Systems

In these systems, a single biometric trait is imaged using multiple sensors in order to extract diverse information from multiple images that may or may not be spatially registered. For example, a system may record the two-dimensional texture content of a person's face using a CCD camera and the three-dimensional surface shape of the face using a range sensor in order to perform authentication. The introduction of a new sensor (in this case, the range sensor) to measure the facial surface variation increases the cost of the multibiometric system. However, the availability of multi-sensor data pertaining to a single trait can assist the *segmentation* and *registration* procedures also besides improving the matching accuracy.

Marcialis and Roli [2] discuss a scheme to fuse the fingerprint information of a user obtained using an optical and a capacitive fingerprint sensor (spatial registration between the two sensors is not necessary in this case). The authors, in their work, indicate that the two sensors provide complementary information thereby providing better matching accuracy. They also suggest the possibility of employing a dynamic sensor selection scheme [3] wherein, based on the nature of the input data obtained from the two sensors, the information from only one of the sensors may be used to perform recognition. Chen et al. [4] examine the face images of an individual obtained using a thermal infrared camera and a visible light camera. They demonstrate that integrating the evidence supplied by these two images (both at the scorelevel and ranklevel) improves matching performance. Socolinsky and Selinger [5] and Heo et al. [6] also demonstrate the benefits of using thermal infrared and visible light imagery for face recognition.

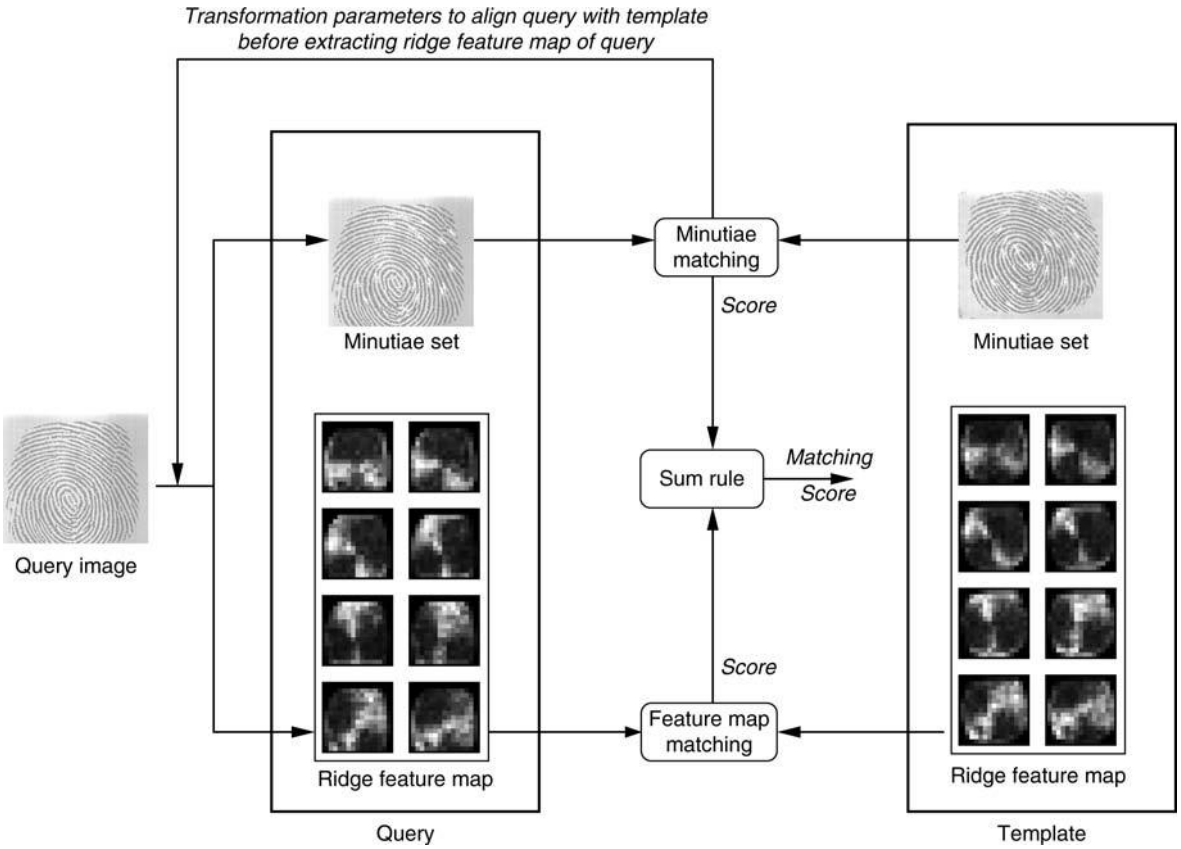
Multi-Algorithm Systems

In these systems, the same biometric data is processed using multiple algorithms. For example, a texture-based algorithm and a minutiae-based algorithm can operate on the same fingerprint image in order to extract diverse feature sets that can improve the performance of the system [7]. This does not require the use of new sensors and, hence, is cost-effective. Furthermore, the user is not required to interact with multiple sensors thereby enhancing user convenience. However, it does require the introduction of new feature extractor and/or matcher modules, which may increase the computational requirements of the system (Fig. 2).

A multi-algorithm system can use multiple feature sets (i.e., multiple representations) extracted from the same biometric data or multiple matching schemes operating on a single feature set. Lu et al. [8] discuss a face recognition system that employs three different feature extraction schemes (principal component analysis (PCA), independent component analysis (ICA), and linear discriminant analysis (LDA)) to encode (i.e., represent) a single face image. The authors postulate that the use of different feature sets makes the system robust to a variety of intra-class variations normally associated with the face biometric. Experimental results indicate that combining multiple face classifiers can enhance the identification rate of the biometric system. Han and Bhanu [9] present a context-based gait recognition system which invokes and combines two gait recognition classifiers based on the walking surface. A probabilistic approach is used to combine the participating classifiers. The authors demonstrate that using context information in a fusion framework has the potential to improve the identification rate of the system. Jain et al. [10] fuse the evidence of three different fingerprint matchers to determine the similarity between two minutiae sets. The three minutiae matchers considered in their system are based on the Hough transform, one-dimensional string matching, and two-dimensional dynamic programming. They observe that the matching performance obtained by combining two of the three matchers is comparable to combining all the three matchers. Factors such as the correlation between component algorithms, the disparity in their matching accuracies, and the fusion methodology adopted significantly impact the performance obtained after fusion.

Multi-Instance Systems

These systems use multiple instances of the same body trait and are also referred to as multi-unit systems in the literature. For example, the left and right index fingers, or the left and right irises of an individual may be used to verify an individual's identity. These systems generally do not necessitate the introduction of new sensors nor do they entail the development of new feature extraction and matching algorithms and



Sources of Information in Biometric Fusion. **Figure 2** The multi-algorithm fingerprint matcher designed in [7]. The system utilizes both minutiae and texture information to represent and match two fingerprint images (query and template). The minutiae-matching module provides the transformation parameters necessary to align the query image with the template before extracting the texture information from the former. The texture information is represented using ridge feature maps.

are, therefore, cost-effective. However, in some cases, a new sensor arrangement might be necessary in order to facilitate the simultaneous capture of the various units/instances. Automated fingerprint identification systems (AFIS) that obtain ten-print information from a subject can benefit from sensors that are able to rapidly acquire impressions of all ten fingers. Multi-instance systems are especially beneficial to users whose biometric traits cannot be reliably captured due to inherent problems. For example, a single finger may not be a sufficient discriminator for a person having dry skin. However, the integration of evidence across multiple fingers may serve as a good discriminator in this case. Similarly, an iris system may not be able to image significant portions of a person's iris due to drooping eyelids. The consideration of both the irides will result in the availability of more texture information that can be used to establish the individual's identity in a more reliable manner. Multi-instance systems are often necessary in applications where the size of the system database (i.e., the number of enrolled individuals) is very large (FBI's database currently has ~50 million ten-print images, and multiple fingers provide additional discriminatory information).

Multi-Sample Systems

A single sensor may be used to acquire multiple samples of the same biometric trait in order to account for the variations that can occur in the trait, or to obtain a more complete representation of the underlying trait. A face system, for example, may capture (and store) the frontal profile of a person's face along with the left and right

profiles in order to account for variations in the facial pose. Similarly, a fingerprint system equipped with a small-sized sensor may acquire multiple dab prints of an individual's finger in order to obtain images of various regions of the fingerprint. A mosaicing scheme may then be used to stitch the multiple impressions and create a composite image. One of the key issues in a multi-sample system is to determine the *number* of samples that has to be acquired from an individual. It is important that the procured samples represent the *variability* as well as the *typicality* of the individual's biometric data. To this end, the desired relationship between the samples has to be established beforehand in order to optimize the benefits of the integration strategy. For example, a face recognition system utilizing both the frontal- and side-profile images of an individual may stipulate that the side-profile image should be a three-quarter view of the face [11]. Alternately, given a set of biometric samples, the system should be able to automatically select the "optimal" subset that would best represent the individual's variability.

Multimodal Systems

These systems combine the evidence presented by different body traits to establish an identity. For example, some of the earliest multimodal biometric systems utilized face and voice features to establish the identity of an individual [12]. Physically uncorrelated traits (e.g., fingerprint and iris) are expected to result in better *improvement* in performance than the correlated traits (e.g., voice and lip movement). The cost of deploying these systems is substantially more due to the requirement of new sensors and, consequently, the development of appropriate user interfaces. The identification accuracy can be significantly improved by utilizing an increasing number of traits although the *curse-of-dimensionality* phenomenon would impose a bound on this number. The curse-of-dimensionality limits the number of attributes (or features) used in a pattern classification system when only a small number of training samples is available. The number of traits used in a specific application will also be restricted by practical considerations such as the cost of deployment, enrollment time, throughput time, expected error rate, user habituation issues, etc.

Hybrid Systems

Chang et al. [13] use the term *hybrid* to refer to systems that integrate a subset of the five scenarios discussed earlier. For example, Brunelli and Falavigna [12] describe an arrangement in which two speaker recognition algorithms are combined with three face recognition algorithms at the match score and rank levels via a HyperBF network. Thus, the system is multi-algorithmic as well as multimodal in its design. Hybrid systems attempt to extract as much information as possible from the various biometric modalities.

Summary

Multibiometric systems can consolidate different pieces of evidence. Besides the above scenarios, it is also possible to use biometric traits in conjunction with nonbiometric identity tokens in order to enhance the authentication performance. For example, Jin et al. [14] discuss a dual factor authenticator that combines a pseudorandom number (present in a token) with a facial feature set in order to produce a set of user-specific compact codes known as BioCode. The pseudorandom number and the facial feature sets are fixed in length, and an iterated inner product is used to generate the BioCode. When an individual's biometric information is suspected to be compromised, the token containing the random data is replaced, thereby revoking the previous authenticator. The use of biometric and nonbiometric authenticators in tandem is a powerful way of enhancing security. However, some of the inconveniences associated with traditional authenticators remain (such as "Where did I leave my token?").

Another category of multibiometric systems combine primary biometric identifiers (such as face and fingerprint) with soft biometric attributes (such as gender, height, weight, eye color, etc.). Soft biometric traits cannot be used to distinguish individuals reliably since the same attribute is likely to be shared by several different people in the target population. However, when used in conjunction with primary biometric traits, the performance of the authentication system can be significantly enhanced [15]. Soft biometric attributes also help in filtering (or indexing) large biometric databases by limiting the number of entries to be searched in the database. For example, if it is determined (automatically or manually) that the

subject is an "Asian Male," then the system can constrain its search to only those identities in the database labeled with these attributes. Alternately, soft biometric traits can be used in surveillance applications to decide if at all primary biometric information has to be acquired from a certain individual. Automated techniques are to estimate soft biometric characteristics which is an ongoing area of research and is likely to benefit law enforcement and border control biometric applications.

Related Entries

- ▶ Multibiometrics
- ▶ Multiple Classifier Systems
- ▶ Multispectral and Hyperspectral Biometrics
- ▶ Soft Biometrics

References

1. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics, 1st edn. Springer, New York, USA (2006)
2. Marcialis, G.L., Roli, F.: Fingerprint verification by fusion of optical and capacitive sensors. Pattern Recognit. Lett. **25**, 1315–1322 (2004)
3. Giacinto, G., Roli, F.: Dynamic classifier selection based on multiple classifier behaviour. Pattern Recognit **34**, 1879–1881 (2001)
4. Chen, X., Flynn, P.J., Bowyer, K.W.: IR and visible light face recognition. Comput. Vision Image Understand. **99**, 332–358 (2005)
5. Socolinsky, D.A., Selinger, A.: Thermal face recognition over time. In: Proceedings of the Seventeenth International Conference on Pattern Recognition (ICPR), vol. 4., pp. 187–190 (2004)
6. Heo, J., Kong, S., Abidi, B., Abidi, M.: Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In: IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum, Washington D.C., USA, pp. 94–99 (2004)
7. Ross, A., Jain, A.K., Reisman, J.: A hybrid fingerprint matcher. Pattern Recognit. **36**, 1661–1673 (2003)
8. Lu, X., Wang, Y., Jain, A.K.: Combining classifiers for face recognition. In: IEEE International Conference on Multimedia and Expo (ICME), vol. 3., Baltimore, USA, pp. 13–16 (2003)
9. Han, J., Bhanu, B.: Gait recognition by combining classifiers based on environmental contexts. In: Proceedings of Fifth International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), Rye Brook, USA, pp. 416–425 (2005)
10. Jain, A.K., Prabhakar, S., Chen, S.: Combining multiple matchers for a high security fingerprint verification system. Pattern Recognit. Lett. **20**, 1371–1379 (1999)
11. Hill, H., Schyns, P.G., Akamatsu, S.: Information and viewpoint dependence in face recognition. Cognition **62**, 201–222 (1997)
12. Brunelli, R., Falavigna, D.: Person identification using multiple cues. IEEE Trans. Pattern Anal. Machine Intell. **17**, 955–966 (1995)
13. Chang, K.I., Bowyer, K.W., Flynn, P.J.: An evaluation of multimodal 2D+3D face biometrics. IEEE Trans. Pattern Anal. Machine Intell. **27**, 619–624 (2005)
14. Jin, A.T.B., Ling, D.N.C., Goh, A.: An integrated dual factor authenticator based on the face data and tokenised random number. In: First International Conference on Biometric Authentication, Hong Kong, China, pp. 117–123 (2004)
15. Jain, A.K., Nandakumar, K., Lu, X., Park, U.: Integrating faces, fingerprints and soft biometric traits for user recognition. In: 1Proceedings of ECCV International Workshop on Biometric Authentication (BioAW), vol. LNCS 3087, Prague, Czech Republic, pp. 259–269. Springer, Berlin (2004)

Speaker Authentication

▶ Speaker Recognition, Standardization

Speaker Biometrics

▶ Speaker Recognition, Standardization

Speaker Change Detection

▶ Speaker Segmentation

Speaker Classification

Speaker Classification is a technology that uses information from the stream of speech to place the speaker into a category such as female versus male, young versus old, native versus non-native speaker.

▶ Speaker Recognition, Standardization

Speaker Clustering

The *clustering* process, in general, can be defined as an unsupervised classification of data, i.e., without any a priori knowledge about the classes or the number of classes. In our task, the clustering process should result, ideally, in a single cluster for every speaker identity.

► Speaker Segmentation

Speaker Databases and Evaluation

Alvin F. Martin
National Institute of Standards and Technology
Gaithersburg, Maryland, USA

Introduction

Expanding interest in the use of biometrics for security purposes has brought increasing attention to the use of speech as a biometric. Speech fits naturally into the list of likely biometric modalities. It is an activity engaged in by essentially everyone, and is one of the primary means by which people identify those whom they know.

But speaker recognition has not heretofore been seen as among the most useful biometrics for general security applications. There has been a much more developmental effort on the use of face, fingerprint, and iris. Recognition of speakers by voice has been seen as more of a niche application, largely because of the special difficulties associated with the collection of quality speech input, and perhaps because of a particular advantage may offer.

This introduction briefly discusses some key issues related to speaker recognition as a biometric. In the following section some of the main databases that have been used for speaker recognition research and evaluation are discussed. The final section deals with the leading technology evaluations of speaker recognition that have been conducted and are ongoing.

Speaker recognition may be divided into speaker identification (many to one decision) and speaker verification or speaker detection (one to one decision). Perhaps because of the performance limitations or because of the difficulty of collecting very large (in the tens of thousands or more) speaker databases, the research community has in recent years focused on the latter. This represents the areas of current practical applications better, and of course ultimately superior performance for the latter would make the former possible.

Defining a "standard" test for speech matching is not simple. Numerous environmental factors affect the quality of any voice signal collected, and these may, depending on the collection configuration and circumstances, be very difficult to control. And there are many choices of protocol to be made, involving in particular the type of speech and specific words, as well as the amount of speech, to be collected. These issues are very much application dependent and operating consensus is very hard to achieve.

Best performance in voice recognition is achieved when a consistent wideband high quality audio channel is available for all speech input. But the needed quiet room environment can be expensive and often impractical to set up, and may be rather demanding on the user in terms of speaking into a close talking microphone. Meanwhile, competing biometrics may more easily provide similar capability.

The particular advantage offered by voice as a biometric is that it is transmissible over telephone channels, and telephone handsets, landline or cellular, are ubiquitous in modern society. The variability of telephone handsets and telephone channels makes the recognition task far more difficult and degrades the quality of performance. Nevertheless this has been the area of greatest application interest, and thus of greatest interest for evaluation.

One key distinction among speaker recognition applications is the type of involvement of the speaker in the process. The speaker may or may not be aware of the recognition process, and if aware, may or may not seek to cooperate with it.

Applications involving access, whether to a physical location or to information, are likely to involve cooperative and motivated users. The system can then prompt the speaker to say a specific phrase, or even a previously agreed upon passphrase (perhaps an account number), allowing the recognition to be text-dependent and even combined with a pin number for greater effective performance. Commercial applications often rely on the use of short phrases spoken by cooperative users, with the system's knowledge of what is to be said (text-dependence) helping to aid performance despite the

limited amount of speech involved and the difficulties posed by the variable conditions of the telephone channels.

Forensic applications, on the other hand, will involve either an unaware or uncooperative user, and other applications will involve listening in on unaware speakers. Here text-dependent recognition is not an option. A possible advantage of this type of application, however, is that it may be possible to collect rather long durations of speech from the speakers, whereas a cooperative scenario requires that valid speakers be able to enroll and obtain access after brief speaking intervals. This can allow systems to learn more about a target speaker's speaking style and idiosyncrasies. The frequency of occurrence of particular words and phrases in someone's natural (determined with the aid of automatic speech recognition technology for word transcription) may powerfully aid recognition performance.

Databases

The era of standard corpora (or databases) for speech processing applications began in the mid-1980's as modest priced computers became capable of performing the necessary signal processing and the costs of storage media fell significantly. The Speech Group at NIST (National Institute of Standards and Technology) played an early role in making the corpora of interest available at reasonable cost in CD-ROM format. Since its founding in 1992, the Linguistic Data Consortium (LDC) at the University of Pennsylvania has been the primary repository of speech corpora in the United States. (ELRA, the European Language Resources Association, plays a similar role in Europe.) The corpora described here are available through the LDC and are described in its online catalog (www.ldc.upenn.edu/catalog).

There are particular properties of the corpora which are needed to support speaker recognition research. A substantial number of different speakers must be included, and most particularly, there need to be a number of different recorded sessions of each speaker. Applications require speakers to enroll in the system at one time and to be successfully detected at a later time. Particularly when recorded over time, multiple recording sessions with varying telephone channels are essential to represent this. Moreover, telephone handsets vary, so it is desirable, for most real-world applications, to have different sessions using different handsets. It has been seen that recognition performance over the telephone is considerably better if speakers can use the same handset during both the training (enrollment) and the test. This is particularly so if impostor speakers use different handsets from speakers of interest, as is typically inherently the case in most collection protocols. But this results in doing channel recognition rather than speaker recognition. Thus a corpus such as Macrophone (the U.S. contribution to the international Polyphone corpus), collected to support multiple types of speech research and containing telephone speech of a variety of types from a large number of speakers, has been of limited usefulness for speaker recognition because of having only a single session for each speaker.

One early corpus widely used for speaker research was TIMIT, produced from a joint effort by Texas Instruments (TI) and the Massachusetts Institute of Technology (MIT), along with the Stanford Research Institute (SRI) with sponsorship by DARPA (Defense Advanced Research Projects Agency). TIMIT is a corpus of read speech, containing 10 phonetically diverse sentences spoken by each of 630 speakers chosen to represent 8 major dialect regions of the United States. Its basic implementation consists of high quality microphone speech, but versions of the data sent through a lower quality microphone channel or different types of telephone channels were also produced.

TIMIT was collected for multiple types of speech processing, but was very popular in speaker identification/recognition research through much of the 1990's, partly because few alternatives were widely available and partly because its limited vocabulary and high recording quality supported the attainment of impressive text-dependent performance results. It was a source of some frustration to leading researchers at speaker recognition workshops held in the 1990s that paper after paper discussed systems performance on TIMIT, rather than on any "real" data.

An early corpus collected specifically for speaker recognition was the KING Corpus. It involved 51 male speakers from whom ten sessions of about 30 seconds each were collected. The speech was collected simultaneously over a wideband channel and a narrowband telephone channel. There were 25 speakers whose speech was collected in New Jersey, and 26 whose speech was collected on in San Diego. For the San Diego speakers,

researchers attempting to do speaker detection noted that there was a "great divide" between the first five and second five of the ten sessions involving narrowband speech. The spectral slope characteristics turned out to be very different on the two sides of this divide. Much effort was devoted to understanding and coping with this phenomenon, and this led to greater awareness of the effects of channel characteristics for speaker recognition using telephone speech, and considerable later research effort to compensate for such channel differences.

A third early corpus for text-dependent recognition of high quality speech was known as the YOHO Corpus. It was collected (like KING) by ITT under a US government contract in 1989. There were 138 speakers each of whom had 4 enrollment and ten verification sessions. Each session involved speaking "combination locks" each consisting of three two digit numbers. There were 24 spoken phrases in the enrollment sessions, and 4 in the verification sessions. This was clearly intended for access applications involving cooperative speakers.

It does not appear that these early corpora were used in multi-site evaluation, but were used extensively in evaluating individual site research projects. As will be noted, it has been difficult to find sufficient interest and agreement on protocols for text-dependent evaluation in the speaker arena. Table 1 summarizes these early corpora.

The modern era in the collection of corpora for speaker recognition, perhaps, began with the collection of the Switchboard Corpus for DARPA by TI in the early 1990s. This collection of about 2,400 two-sided telephone conversations from over 500 participating speakers was originally intended for multiple purposes, including word spotting, topic spotting, and speaker spotting in the terminology used at the time. An automatic system was created which allowed registered participants to call in at specified times to a "robot operator" which attempted to contact other registered participants and initiate a two-way conversation on one of about 70 pre-specified topics that the participants had indicated would be acceptable. Thus the conversants generally engaged in an at least somewhat serious discussion for five minutes or more with someone whom they did not know. A speaker's topic and conversational partner were in general never repeated in different conversations. A subset of the participating speakers was encouraged to make a sizable (double-digit) number of different conversations and to use multiple telephone handsets in them.

Switchboard-1 (so denoted when similar corpora followed) was used in a couple of limited U.S. government sponsored evaluation of speaker spotting (and a similar evaluation of topic spotting) in the early 1990s, and it proved to be a popular corpus for further study and research. Somewhat surprisingly it was used in subsequent years for general evaluation of automatic speech (word) recognition, as the focus of such evaluation shifted to natural unconstrained conversational speech. And it 1996 it provided the data for the first of the series of ► NIST (SRE's) Speaker Recognition Evaluations discussed below. A subset of 40 of the most prolific corpus speakers was used as the target speaker set in this evaluation.

The success of Switchboard-1, particularly for speaker recognition, led to the collection of the multi-part Switchboard-2 and Switchboard Cellular Corpora. Each involved hundreds of speakers taking part in a number of different conversations using multiple telephone handsets. This was important as the early NIST evaluations established that telephone handset variation between training and test affected system performance very much, and the desire was to truly recognize speakers and not merely handsets.

Speaker Databases and Evaluation. Table 1 Some early corpora used for speaker recognition

Year	Corpus	Size	Types of speech
Early 1980s	TIMIT	630 speakers of eight major US English dialects, 10 sentences each; alternative versions run original wideband data through other specified channels	Read speech of phonetically rich sentences
1987	KING	51 male speakers (25 New Jersey, 26 San Diego), 10 sessions each recorded on both a wide-band and a narrow-band channel	Sessions contain 30 s of speech on an assigned topic
1989	YOHO	138 speakers with 4 enrollment sessions (24 phrases) and 10 test sessions (4 phrases)	"Combination lock" phrases

Speaker Databases and Evaluation. **Table 2** The Switchboard corpora; Collection Years are Approximate

Year	Corpus	Size	Types of speech
1990/1991	SWBD I	543 speakers, 2400 two-sided conversations	USA conversational telephone speech on assigned topics
1996	SWBD II phase 1	657 speakers, 3638 conversations	Primarily US Mid-Atlantic, conversational telephone
1997	SWBD II phase 2	679 speakers, 4472 conversations	Primarily US Mid-West, conversational telephone
1997/1998	SWBD II phase 3	640 speakers, 2728 conversations	Primarily US South, conversational telephone
1999/2000	SWBD cellular p1	254 speakers, 1309 conversations	Primarily cellular GSM, USA conversational
2000	SWBD cellular p2	419 speakers, 2020 conversations	Cellular, largely CDMA, USA conversational

The Switchboard-2 Corpora each concentrated largely on speakers from a specific area of the United States, relying mainly on college students or early post-college age people. Switchboard Cellular was collected in the light of the increasing use of cellular telephone handsets in the United States.

The Switchboard Corpora supplied the bulk of the evaluation data used for the annual NIST evaluations from 1996 to 2003. Table 2 summarizes these corpora.

Around 2003 the LDC moved to a somewhat different collection model from that used in the Switchboard Corpora. The "Fisher" platform was similar to that used for Switchboard, but it could also initiate a search for paired conversants without one party initiating matters with a call into the system. It was to prove useful in new corpus collections for general speech recognition and for language recognition, but was also applied to the speaker recognition collection. For this purpose the multi-part Mixer Corpus has been collected. It was used in the 2004, 2005, and 2006 SRE's, and will be used in the 2008 SRE.

The Mixer collections have expanded the types of speaker data collected in two major ways. The first is the inclusion of conversations in multiple languages. LDC recruited a sizable number of bilingual speakers (with English as one language) and utilized the collection protocol to pair up speakers of a non-English language, who received a bonus for talking in their other language. It became feasible, for example, to have certain specified days devoted to the collection of calls in specified languages. This supported investigation of the effect of language, and of language change between training and test, in speaker recognition performance.

In addition, the Mixer corpora have included some conversations in which participants were recorded simultaneously over the telephone and over eight or more different microphones. These included a range of close talking, near-field, and far-field microphones to support comparison of performance with the different types, and of cross channel conditions between training and test. This was accomplished by having select groups of participants come to a special room at two collection sites where all of the microphones could be carefully placed while they used the cell phones provided to call the automatic system and be paired with participants in the usual way.

The Mixer 5 Corpus collected in 2007 contains a further variation on this theme. Its 300 speakers each participated in a series of six structured "interviews" of about a half hour each, occurring over at least three different days. The bulk of each interview involves conversational speech, but with an interviewer who is present in the room and provides appropriate prompts. The subject's speech is recorded over a dozen or so carefully placed microphones, but not over a telephone line. Over the course of the six sessions the subject gets to know the interviewer, and this changes the nature of the spoken dialog. Each interview also contains a brief period of standard questions repeated, and periods of different types of read speech. Each participant also makes two simulated phone calls where side tones are used to encourage each, rather high or rather low, vocal effort. Each interview subject is also paired in the usual way in about ten regular phone conversations with unknown interlocutors outside of the interviews. This data will be used in the upcoming NIST

Speaker Databases and Evaluation. **Table 3** The Mixer corpora; collection years are approximate

Year	Corpus	Size	Types of speech
2003	MIXER p1 and p2	600 speakers with 10 or more calls 200 with 4 cross-channel calls	Conversational, some calls in four non-English languages
2005	MIXER p3	1,867 speakers with 15 or more calls	Conversational, includes calls in 19 languages
2007	MIXER p4	200 speakers making 10 calls including 4 cross-channel	Conversational, primarily English
2007	MIXER p5	300 speakers doing 6 interviews and generally 10 phone calls	Conversational in interview setting, some read speech

SRE's and may offer some interesting contrasts with previous results.

The Mixer Corpora are discussed further in [1–3]. Table 3 summarizes the Mixer Corpora.

Evaluations

Evaluations of speaker recognition require a sponsor or sponsors and participants. Sponsors must be willing to commit the necessary resources to support an evaluation infrastructure. Most important, they must support the collection of speech databases appropriate to speaker recognition evaluation needs.

Participants must be willing to take part in evaluation, to discuss the systems they develop, and to have their performance results presented to the evaluation community. They must be ready to do this not knowing in advance whether their evaluation performance will compare favorably or unfavorably with that of the other participants.

The most notable series of evaluations of recent years have been those coordinated by the National Institute of Standards and Technology (NIST), an agency of the U.S. Department of Commerce, in Gaithersburg, Maryland, USA. The NIST evaluations have received sponsorship support and guidance from interested U.S. government agencies involved in defense, intelligence and law enforcement.

There were a couple of preliminary evaluations held in 1992 and 1995, each utilizing a limited number of target speakers from the Switcboard-1 Corpus. They did not involve the scoring metric of the later evaluations, described below, and looked at the range of operating points (receiver operating characteristic curves) of each target speaker separately rather than combining them based on a required calibration threshold into a single curve as will be described below. The 1995 evaluation was the first to analyze and note the effect on performance of having a speaker's training and test segments come from the same or different telephone numbers, and thus the same or different telephone handsets. These evaluations each had only about a half dozen participants, mainly from the United States.

The NIST evaluations assumed basically their present form in 1996, and were conducted annually from 1996 to 2006, with the next one set to occur in 2008. These have all included as the core task text-independent speaker detection in the context of conversational telephone speech. The 1996 evaluation selected 40 of the more prolific Switchboard-1 speakers as target talkers, and used other corpus speakers for non-target trials. The subsequent evaluations have all utilized hundreds of speakers from the LDC corpora involved (Switchboard through 2003, Mixer subsequently), and have followed the practice of allowing the target speakers to also serve as impostor speakers for non-target trials. The evaluation plan documents and other information related to these evaluations may be found at http://www.nist.gov/speech/tests/sre/index.html.

Participation in the NIST speaker recognition evaluations has grown steadily and has become worldwide in scope. The number of participating sites has grown to reach approximately 35 in 2006. The number of participants noticeably increased in 2002 and subsequent years, perhaps because of a growing interest in biometric technologies after the events of 2001.

Of the growing number of participants in recent years, only about half a dozen have been sites in the United States, with a majority in Europe, and an increasing number from the Far East. The greatest numbers of participants have been from the U.S.,

France, and China. Other participants have been from Canada, various European countries, Singapore, Australia, Israel, and South Africa.

Most of the sites participating in the NIST evaluations have been from academic institutions. Some government funded research institutions or companies involved in government research have also participated. Not frequently represented, however, have been smaller commercially oriented companies. This may be due in part to the text-independent and research oriented type of evaluation being conducted, but also bespeaks a reticence to participate in evaluations where competitors may show superior performance results.

Evaluation requires a performance measure. For detection tasks there are inherently two types of error. There are trials where the target is present (target trials) but a "false" decision is made by a system. Such errors are misses. And there are trials where the target is not present (non-target or impostor trials) but a "true" decision is made. These are referred to as false alarms. Thus it is possible to speak of a miss rate for target trials and a false alarm rate for non-target trials.

The NIST evaluations have used a linear combination of these two error rates as its primary evaluation metric. A decision cost function (DCF) is defined as

$$DCF = C_{Miss} \times P_{Miss|Target} \times P_{Target} + C_{FalseAlarm} \times P_{FalseAlarm|NonTarget} \times (1 - P_{Target})$$

where C_{Miss} represents the cost of a miss, C_{FA} the cost of a false alarm, and P_{Target} the prior probability of a target trial. These are three somewhat arbitrary and certainly application dependent parameters. The NIST evaluations have used parameter values as hereunder.

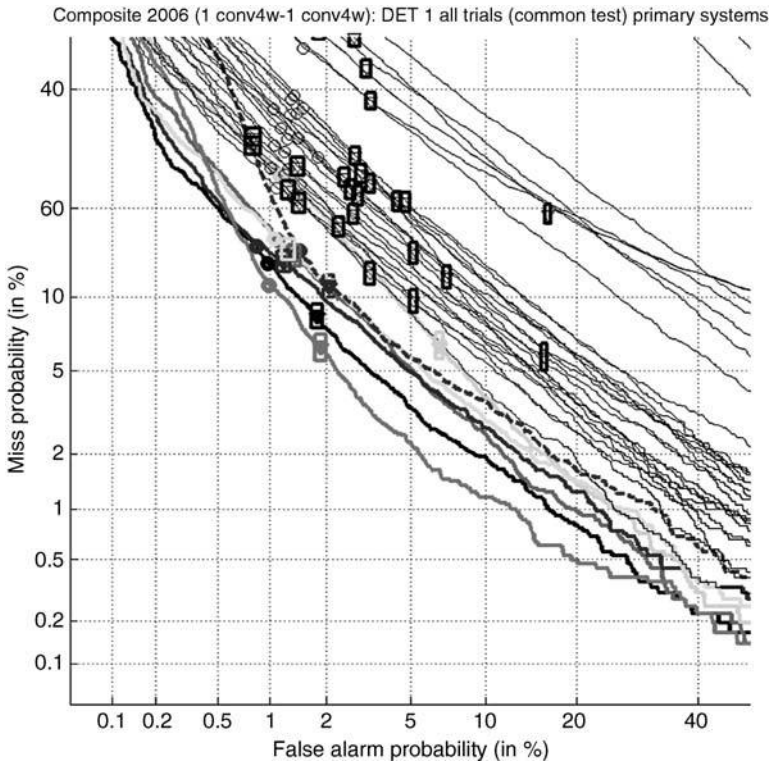
C_{Miss}	$C_{FalseAlarm}$	P_{Target}
10	1	0.01

These have been viewed as reasonable parameters for applications involving an unaware user, where most speech segments examined are likely to be of someone other than the target of interest, but where detecting instances of the target have considerable value. Note that P_{Target} need not represent the actual target richness of the evaluation trials, but may be chosen based on possible applications of interest. The NIST evaluations have generally had an approximately ten to one ratio of non-target to target trials, to minimize the variance of the metric in the light of the parameter values chosen.

A detection task inherently involves two types of error, and a system may be expected to be able to tune its performance to vary the relative frequency of the two error types. In the NIST evaluations, systems have been required to produce not only a decision, but also a score for each trial, where higher scores indicate greater likelihood that the correct decision is "true". A decision threshold may then be varied based on this score to show different possible operating points or tradeoffs between the two types of error. Note that the evaluations have required that this threshold be the same for all target speakers.

The most informative way of presenting system performance in the NIST SRE's has been to draw a curve showing the operating points and the tradeoff in the error rates. This is easily done by varying the decision threshold based on the scores provided. A simple linear plot is know as an ROC (Receiver Operator Characteristic) curve, but a clearer presentation is obtained by putting both error rates on a normal deviate scale to produce what NIST has denoted a DET (Detection Error Tradeoff) curve [4]. This has the nice property that if the underlying error distributions for the miss and false alarm rates are normal, the resulting curve is linear.

Figure 1 shows ▶ DET curves for the systems in the core test done in the 2006 NIST SRE. These are curves representing the performance of the primary systems submitted by over 30 sites participating in the evaluation. Better systems have performance curves closer to the lower left corner of the plot. The actual decision point of each performance curve is denoted by a triangle, and a 95% confidence box is drawn around these, while circles are used to denote the points corresponding to the minimum DCF operating points. The closer these two specially denoted points on each curve, the better the system did at calibrating its decision threshold for hard decisions. For example, for the best performing system shown, the actual decision point has a false alarm rate of about 2% and a miss rate of about 7%, while the minimum DCF point has a false alarm rate of about 1% and a miss rate of about 11%. This gives a sense of the level of current state-of-the-art performance for speaker detection on this type of telephone data.



Speaker Databases and Evaluation. **Figure 1** DET (Detection Error Tradeoff) Curves for the primary systems of participating sites on the core test of the 2006 NIST SRE.

A possible alternative non-parametric information theoretic type of metric has been proposed to be applicable to a range of applications, and has been included as an alternative measure in the most recent NIST evaluations, provided the system specifies that its likelihood scores may be viewed as log likelihood ratios. This metric is discussed in [5].

While the basic detection task has remained fixed, there have been multiple test conditions in most of the evaluations, and these conditions have varied over the years. In particular there has been variation in the durations of the training and test segments. While the earlier evaluations focused on landline phones and the varying types of telephone handsets (carbon-button vs. electrets microphone), in the new millennium there is was greater focus on the effect of cellular transmission and newer types of handsets as these became common in the U.S. Certain additional data sources, such as a small FBI forensic database and a Castilian Spanish corpus known as AHUMADA (both apparently not currently easily available) were used in one or two evaluations.

The earlier evaluations used fixed durations of speech, as determined by an automatic speech detector. Later evaluations allowed more variation in duration within each test condition. Starting in 2001 there was greater interest in longer durations for training and test. This was largely as a result of some research suggesting that with effective word recognition, higher level lexical information about a speaker could be effectively combined with more traditional lower level acoustic information [6]. As a result of the apparent success of such an approach in the 2001 evaluation, a major summer research program was carried out at Johns Hopkins University in the summer of 2001 (see http://www.clsp.jhu.edu/ws2002/groups/supersid/). Since then, "extended" training conditions, where the training consists of multiple (often eight) conversation sets have been a major part of the evaluations. The earlier NIST evaluations are described further in [7–9].

The introduction of Mixer data in 2004 inaugurated a new era in the NIST evaluations. The inclusion of calls in multiple languages and cross language trials introduced a new wrinkle that affected overall

performance. The latest evaluations have also introduced test conditions involving multiple microphones and cross channel trials, that will be a major focus in 2008 and beyond. The recent SRE's are discussed in [10–12].

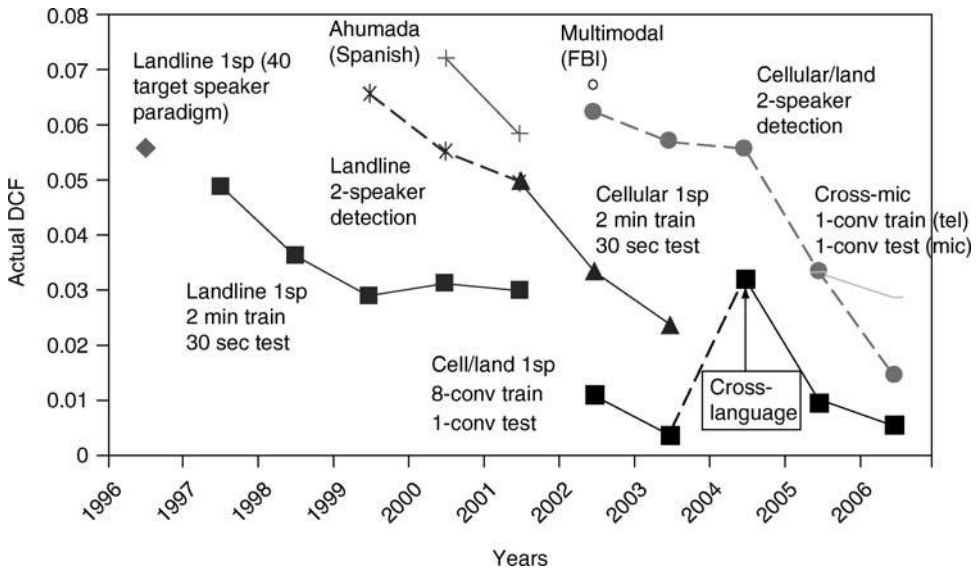
Have the evaluations shown progress in performance capabilities over the years? They have, but changes in the test conditions from year to year and in the types of data used have complicated performance comparisons. Figure 2 from [13] attempts to sort these matters out, and summarizes the DCF scores of the best evaluation systems across ranges of years involving more or less consistent test conditions.

The NIST SRE's have been the most notable evaluations in speaker recognition in recent years. They have concentrated on a basic speaker detection task not tied to any specific current commercial application. This has made it possible for a large range of research sites around the world to participate in these evaluations.

One other notable evaluation in the field was conducted by TNO in the Netherlands in 2003. It featured a protocol very similar to that of the NIST evaluations, but utilized actual forensic data provided by the Dutch police. Its very interesting results are discussed in [14], but the data used was only provided to the evaluation participants for a limited time and purpose and is not otherwise available.

Other efforts have been less successful. Research in speaker recognition technology has been advanced by the series of Odyssey workshops. These were held in Martigny, Switzerland in 1996, Avignon, France in 1998, Crete, Greece in 2001 (where the name "Odyssey" was adopted), Toledo, Spain in 2004, San Juan, Puerto Rico in 2006, and Stellenbosch, South Africa in 2008. For the 2001 workshop an evaluation track was included. This included both a text-independent track based on the preceding NIST evaluation, and a text-dependent track. Participation, particularly in the text-dependent track, was very limited, perhaps demonstrating the difficulty of persuading companies or organizations to participate in this inherently application specific and more immediately commercially oriented field.

The European Union has sponsored a multi-year program to develop biometric technologies denoted as "BioSecure" (http://www.biosecure.info/), with speaker as one of the included technologies. Evaluation is intended to be part of this program, in particular including evaluation of the fusion of multiple biometrics. As of 2007, however, speaker recognition evaluation appears not to have begun.



Speaker Databases and Evaluation. **Figure 2** DCF (Decision Cost Function) values for the best (lowest DCF) systems on different roughly comparable evaluation conditions over multiple years during the course of the NIST SRE's from 1996 to 2006.

The NIST evaluations will resume in 2008, and may be held in alternate years in the future. They will feature an increased emphasis on cross channel recognition. Whereas in 2005 and 2006 the core test involved only telephone speech, with cross channel (train on telephone, test on microphone) as an optional additional test, the core test condition is expected to require processing of a mix of training or test segments including both telephone and microphone speech, with some of the trials including different channels in training and test. This will utilize at least both types of data as in Mixer 3 and Mixer 5. Evaluation performance, however, will be subsequently analyzed to distinguish performance on telephone, microphone, and cross-channel trials. A number of different microphone types from the Mixer 5 data will be included.

Related Entries

► Performance Evaluation, Overview
► Speaker Recognition, Overview

References

1. Cieri, C., Campbell, J.P., Nakasone, H., Miller, D., Walker, K.: The Mixer Corpus of Multilingual, Multichannel Speaker Recognition Data, LREC 2004: Fourth International Conference on Language Resources and Evaluation, Lisbon (2004)
2. Cieri, C., Andrews, W., Campbell, J.P., Doddington, G., Godfrey, J., Huang, S., Liberman, M., Martin, A., Nakasone, H., Przybocki, M., Walker, K.: The Mixer and Transcript Reading Corpora: Resources for Multilingual, Crosschannel Speaker Recognition Research, LREC 2006: Fifth International Conference on Language Resources and Evaluation (2006)
3. Cieri, C., Corson, L., Graff, D., Walker, K.: Resources for New Research Directions in Speaker Recognition: The Mixer 3, 4 and 5 Corpora**,** Interspeech 2007, Antwerp (August 2007)
4. Martin, A.F., et al.: The DET curve in assessment of detection task performance. In: Proceedings of Eurospeech '97, vol. 4, pp. 1899–1903. Rhodes, Greece (September 1997)
5. Brummer, N., du Preez, J.: Application-independent evaluation of speaker detection. Comput. Speech Lang. **20**(2–3), 230–275 (April–July 2006)
6. Doddington, G.: Speaker recognition based on idiolectal differences between speakers. In: Proceedings of Eurospeech '01, vol. 4, pp. 2521–2524. Aalborg, Denmark (September 2001)
7. Martin, A.F., Przybocki, M.A.: The NIST speaker recognition evaluations: 1996–2001. In: Proceedings of 2001: A Speaker Odyssey, pp. 39–43. pp. 39–43. Chainia, Crete, Greece (June 2001)
8. Martin, A.F., Przybocki, M.A., Campbell, J.P.: The NIST speaker recognition evaluation program. In: Wayman, J. (eds.) et al.: Biometric Systems: Technology, Design and Performance Evaluation, Chapter 8, pp. 241–262. ?pp. 241–262. Springer, Berlin (2005)
9. Przybocki, M.A., Martin, A.F.: NIST speaker recognition evaluation chronicles. In: Proceedings of Odyssey 2004: The Speaker and Language Recognition Workshop. Toledo, Spain (2004)
10. Przybocki, M.A., Martin, A.F., Le, A.N.: NIST speaker recognition evaluation chronicles – Part 2. In: Proceedings of Odyssey 2006: The Speaker and Language Recognition Workshop. San Juan, PR (2006)
11. Przybocki, M.A., Martin, A.F., Le, A.N.: NIST speaker recognition evaluations utililizing the mixer corpora – 2004, 2005, 2006. IEEE Trans. Audio Speech Lang. Process. **15**(7), (2007)
12. Martin, A.F.: Evaluations of automatic speaker classification systems. In: Muller, C. (ed.) Speaker Classification I, pp. 313–329. pp. 313–329. Springer, Berlin (2007)
13. Reynolds, D.A.: Keynote talk. In: Proceedings of Odyssey 2008: The Speaker and Language Recognition Workshop. Stellenbosch, South Africa (January 2008)
14. van Leeuwen, D.A., et al.: NIST and NFI-TNO evaluations of automatic speaker recognition. Comput. Speech Lang. **20**(2), 128–158 (2006)

Speaker Detection

Speaker detection means determining whether or not a particular speaker is present in an audio stream. The term multispeaker detection refers to the task of determining whether a particular known speaker is speaking in an audio stream containing speech from multiple speakers.

► Speaker Segmentation

Speaker Diarization

This task consists of segmenting a conversation involving multiple speakers into homogeneous parts, which contain the voice of only one speaker, and grouping together all the segments that correspond to the same speaker.

► Speaker Segmentation

Speaker Features

Daniel Ramos, Javier Gonzalez-Dominguez, Doroteo T. Toledano, Joaquin Gonzalez-Rodríguez
ATVS – Biometric Recognition Group. Escuela Politecnica Superior, Universidad Autonoma de Madrid, Spain

Synonyms

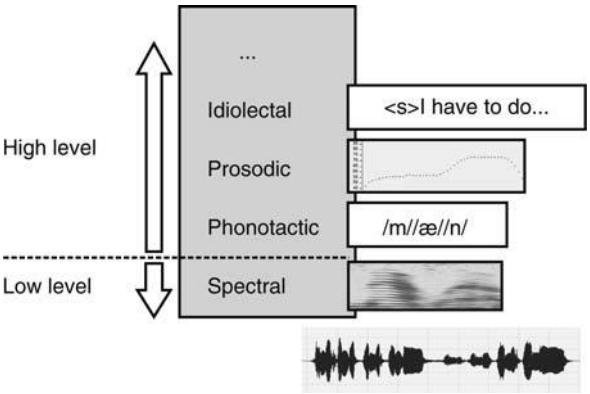
Observations from speech; Speaker parameters

Definition

Speaker features are measurements extracted from the speech signal with the objective of determining the identity of a given speaker. In voice biometrics, speaker features whose source is known are typically used to build ► speaker models. Then, speaker features of unknown source are compared with the enrolled models in order to obtain measures of similarity. The identity of the speaker influences the speech production process in many different ways, due to vocal tract configuration, language spoken, social context, education, etc. Thus, several levels of identity can be identified in the speech signal, e.g., spectral, phonetic, prosodic, etc. Speaker features can be extracted at any of this identity levels, and therefore the speaker recognition process follows in essence a multilevel approach.

Identity Information in the Speech Signal

The ► identity levels in the speech signal are configured by the speech production process, which is the subject of study of phoneticians and other areas such as engineering, physics or signal processing [1, 2, 3, 4]. There are two main stages in voice production: (1) language generation and (2) speech production; and speaker specificities are introduced in both components. In the field of speaker recognition these two components correspond to so-called high-level (linguistic) and low-level (spectral) characteristics. Automatic speaker recognition systems will intend to take advantage of the different sources of information available in the speech signal, combining them in the best possible way for every speaker [5, 6]. Figure 1 illustrates these different identity levels in the speech signal. The information extracted in each of these groups of levels can be summarized as follows:

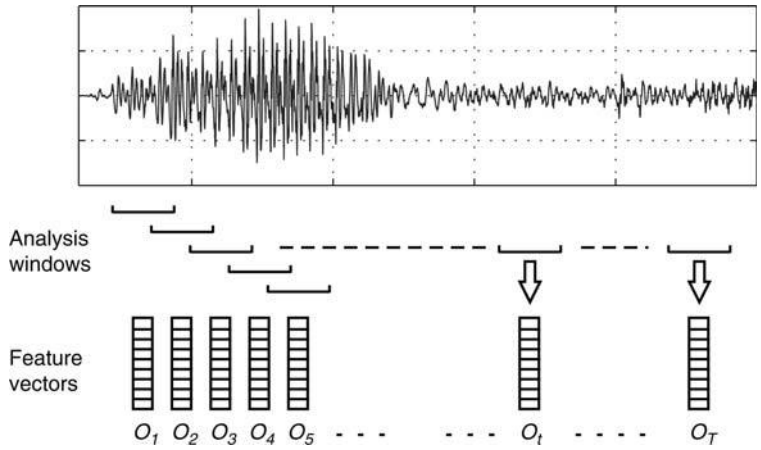


Speaker Features. Figure 1 Identity levels in the speech signal.

- *Spectral level.* The information about the speaker identity is extracted from the spectrum of the speech signal, analyzed in short-time windows. The spectrum of the speech signal is directly related to the dynamic configuration of the vocal tract, which presents speaker-dependent specificities.
- *Higher levels.* Several sublevels can be found here. For instance, at the phonotactic level, the information about the identity of the speaker is embedded in the particular use of the phones and syllables and their realizations. At the prosodic level, parameters like instantaneous energy, intonation, speech rate, and unit durations are analyzed, which are known to be speaker-dependent. At the idiolectal level, the information about speaker identity relies on the particular use of the words and language in general, which depends not only on the speaker, but also on many other sociolinguistic conditions.

Short-Term Spectral Feature Extraction

The analysis at spectral level of the speech signal is based on classic Fourier analysis. However, an exact definition of Fourier transform cannot be directly applied because speech signal cannot be considered



Speaker Features. **Figure 2** Short-term feature extraction.

stationary due to constant changes in the articulatory system within each speech utterance.

To solve these problems, speech signal is split into a sequence of short segments in such a way that each one is short enough to be considered pseudo-stationary. The length of each segment, also called *window* or *frame*, ranges between 10 and 40 ms (in such a short-time period our articulatory system is not able to significantly change). Finally, a feature vector will be extracted from the short-time spectrum in each window. The whole process, known as *short-term analysis*, is depicted in Fig. 2.

Signal representation or coding from short-term spectrum into a feature vector is one of the most important steps in automatic speaker recognition and continues being subject of research. Many different techniques have been proposed in the literature and generally they are based on speech production models or speech perception models. Most widely used techniques in the state of the art are described as follows.

- *Linear Predictive Coding (LPC)* method, introduced in [7], is based on the assumption that a speech sample can be approximated by a linearly weighted summation of a determined number of preceding samples. In time domain, this can be represented as

$$s^*[n] = \sum_{k=0}^{p} a[k]s[n-k]. \quad (1)$$

Here, $s^*[n]$ is the approximation, or *prediction*, of the speech signal, and $a[k]$ are the LPC coefficients calculated to minimize the total square error

$$E = \sum_n e[n]^2, \quad (2)$$

where $e[n]$ is the error between the real signal value $s[n]$ and predicted value $s^*[n]$, defined as

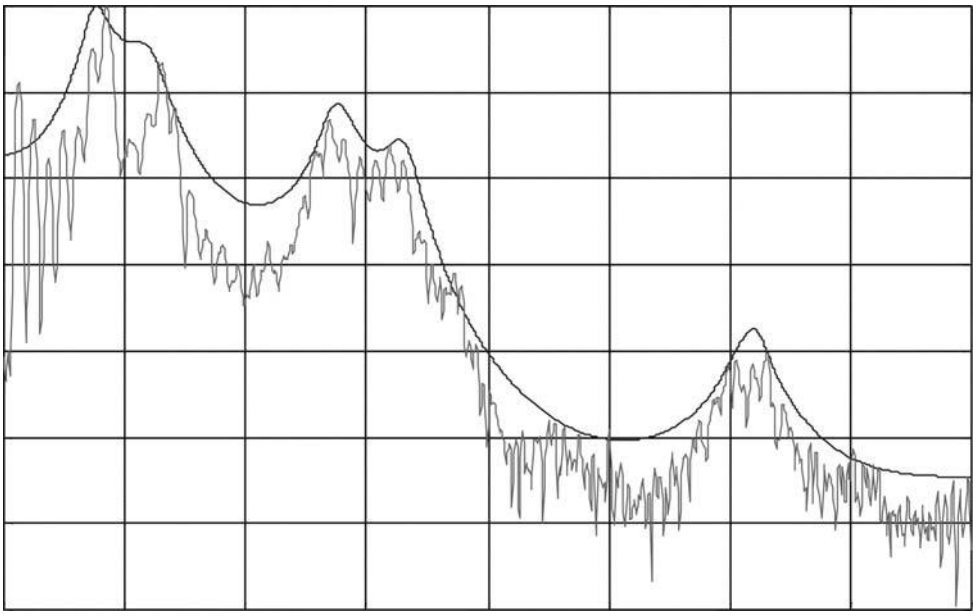
$$e[n] = s[n] - s^*[n] = s[n] - \sum_{k=1}^{p} a[k]s[n-k]. \quad (3)$$

In the domain of the z-transform, $a[k]$ parameters define an all-pole filter $H(z)$, as defined in [1, 7].

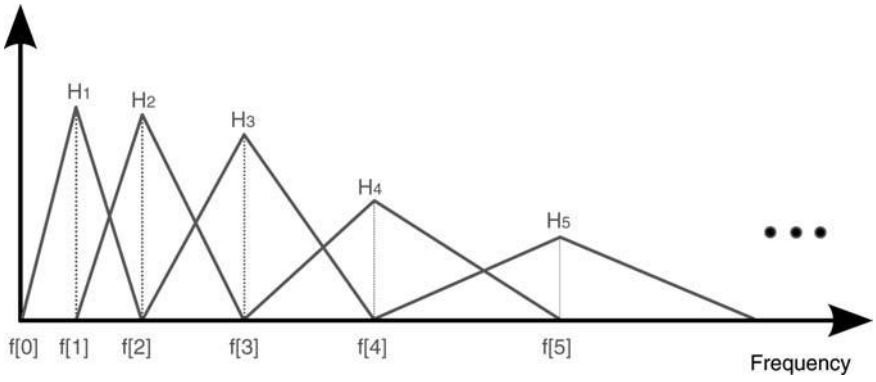
$$H(z) = \frac{1}{1 - \sum_{k=1}^{p} a[k]z^{-k}}. \quad (4)$$

LPC has proved to be a valid way to compress the spectral envelope in an all-pole model with just 10–16 coefficients [1, 3]. Figure 3 shows the representation of the envelope of the short-time spectrum at a given window as modeled by LPC. However, LPC coefficients are strongly correlated among them, which is an undesirable characteristic. Therefore, ▶ cepstrum transform [3, 8] has been proposed in order to obtain pseudo-orthogonal *cepstral* coefficients, yielding Linear Prediction Cepstral Coefficients (LPCC).

- *Mel-Frequency Cepstral Coefficients (MFCC)* proposed in [9] are the most extensively used parameters at the spectral level in automatic speaker recognition systems. The MFCC method first uses a mel-scale filterbank in order to obtain some coefficients from the power spectrum of the speech window. The main aim of mel filtering is to mimic



Speaker Features. Figure 3 LPC modeling of short-term spectrum. The envelope is determined by the LPC filter coefficients defined in Eq. 4.



Speaker Features. Figure 4 Triangular mel-filter bank for typical MFCC feature extraction.

the human hearing behavior by emphasizing lower frequencies and penalizing higher frequencies. Thus, a mel filterbank analyzes the power spectrum using a logarithmic scale. First, a transformation is applied according to the following formula:

$$f_m = 1,125 \times log(1 + f/700), \qquad (5)$$

where f is the linear frequency. Second, a filterbank is applied to the amplitude of the mel-scaled spectrum f_m in order to obtain a vector of outputs from each filter.

Figure 4 shows a typical mel filterbank in the frequency domain. The centers $f[m]$ of the filters $H_m[k]$ are uniformly spaced in the mel scale. Using a DFT of the input signal with N points each filter $H_m[k]$ is given by

$$H_m[k] = \begin{cases} 0 & k < f[m-1] \\ \frac{(k-f[m-1])}{(f[m]-f[m-1])} & f[m-1] \le k \le f[m] \\ \frac{(f[m+1]-k)}{(f[m+1]-f[m])} & f[m] \le k \le f[m] \\ 0 & k > f[m+1], \end{cases} \qquad (6)$$

where $0 < k < N$.

Once filtering is carried out, cepstrum transform is applied to the filter outputs in order to obtain mel frequency cesptrum coefficients.

- *Perceptual Linear Prediction (PLP)* was proposed in [10]. Here, speaker features are calculated in a similar way as LPC coefficients, but previous transformations are carried out in the spectrum of each window aiming at introducing knowledge about human hearing behavior. Details can be found in [10].

As we mentioned earlier, the main aim of the described methods is to extract a feature vector for each frame or window. However, in this independent analysis possible useful information such as coarticulation can be lost. In order to take this kind of information into account, velocity (Δ) and acceleration ($\Delta\Delta$) coefficients are usually obtained from the static window-based information. This Δ and $\Delta\Delta$ coefficients model the speed and acceleration of the variation of cepstral feature vectors across adjacent windows.

High-Level Tokenization

At phonotactic and idiolectal levels, tokenization is the translation from sampled speech into a time-aligned sequence of linguistic units, or *tokens*. Hidden Markov Models (HMM) [11] are widely used for phone, syllable, and word tokenization. HMM as used in speech processing are finite state machines which model the temporal dependency of spectral feature vectors in a probabilistic way [1, 11]. The performance of the tokenization may be improved by the use of language models, which impose some linguistic or grammatical constraints on the high number of combinations of all possible units (phones, syllables or words) [1].

Basic prosodic features as pitch and energy are also obtained at a frame level. The instantaneous pitch can be determined by, e.g., autocorrelation or cepstral-decomposition based methods, usually smoothed with some time filtering [2]. Other important prosodic features are those related to linguistic units duration, speech rate, and all those related to accent. In all those cases, precise segmentation is required [1, 12], i.e., determination of the points in the speech signal where each unit occurs.

Recently, Nonuniform Extraction Region Features (NERF) have been proposed for obtaining high level features [13]. This technique is based on including high-level information in the spectral information at each short-time frame.

Summary

The information about the identity of the speaker extracted from a speech utterance is represented by the speaker features, which can be obtained at different levels. The essay presented the main approaches for speaker feature extraction at the short-time spectral level and at higher levels. The widely used MFCC, LPCC, and PLP features have been described, and several approaches of phonetic and prosodic tokenization have been sketched. Such features will be used to build the models and to compare them with test speech segments in a given voice biometric system.

Related Entries

- ▶ Feature Extraction
- ▶ Speaker Matching
- ▶ Session Effects on Speaker Modeling
- ▶ Speaker Recognition, Overview
- ▶ Speech Analysis

References

1. Huang, X., Acero, A., Hon, H.W.: Spoken Language Processing: A Guide to Theory, Algorithm and System Development. Prentice Hall PTR, Upper Saddle River, NJ (2001)
2. Rabiner, L.R., Schafer, R.W.: Digital Processing of Speech Signals. Prentice Hall, Englewood Cliffs, NJ (1978)
3. Deller, J.R., Hansen, J.H.L., Proakis, J.L., Discrete-Time Processing of Speech Signals, 2nd Ed. Wiley, New York (1999)
4. Gonzalez-Rodriguez, J., Toledano, D.T., Ortega-Garcia, J.: Voice biometrics. In: Anil K. Jain, Patrick Flynn and Arun A. Ross (eds.) Handbook of Biometrics. Springer, Berlin (2007)
5. Reynolds, D.A.: The SuperSID project: Exploiting high-level information for high-accuracy speaker recognition. In: Proceedings of ICASSP (2003)
6. Doddington, G.: Speaker recognition based on idiolectal differences between speakers. In: Proceedings of Eurospeech, pp. 2517–2520. Aalborg, Denmark (2001)
7. Makhoul, J.: Spectral analysis of speech by linear prediction. IEEE Trans. Audio Electroacoust. **21**, 140–148 (1973)
8. Furui, S.: Cepstral analysis technique for automatic speaker verification. IEEE Trans. Acoust. Speech, Signal Processing **29**, 254–272 (1981)
9. Bridle, J.S., Brown, M.D.: An experimental automatic word recognition system. Technical Report 1003, Joint Speech Research Unit, Ruislip, England (1974)
10. Hermansky, H., Hanson, B., Wakita, H.: Perceptually based linear predictive analysis of speech. In: Proceedings of ICASSP, Vol. 10, pp. 509–512. (1985)

11. Rabiner, L.R.: A tutorial on hidden markov models and selected applications in speech recogniton. Proc. IEEE **77**, 257–286 (1989)
12. Toledano, D.T., Hernandez-Gomez, L., Villarrubia-Grande, L.: Automatic phonetic segmentation. IEEE Trans. Speech Audio Process **11**, 617–625 (2003)
13. Kajarekar, S., Ferrer, L., Sonmez, K., Zheng, J., Shriberg, E., Stolcke, A.: Modeling NERFs for speaker recognition In: Proceedings of Odyssey, pp. 51–56 Toledo, Spain (2004)

Speaker Identification and Verification, SIV

► Speaker Recognition, Standardization

Speaker Indexing

The process of determining, who is talking when, is an integral element of speech data monitoring and content-based data mining applications.

► Speaker Segmentation

Speaker Matching

Jean-François Bonastre, Driss Matrouf
LIA - CERI University of Avignon, Avignon, France

Synonyms

Speaker recognition engine; Voice biometric engine

Definition

Speaker matching aims to compare the acquired data corresponding to an individual against the template feature set stored in the database. Depending on the operating mode, the comparison could be done using only the template related to a given person (detection or verification tasks) or with all the templates of the database (identification task).

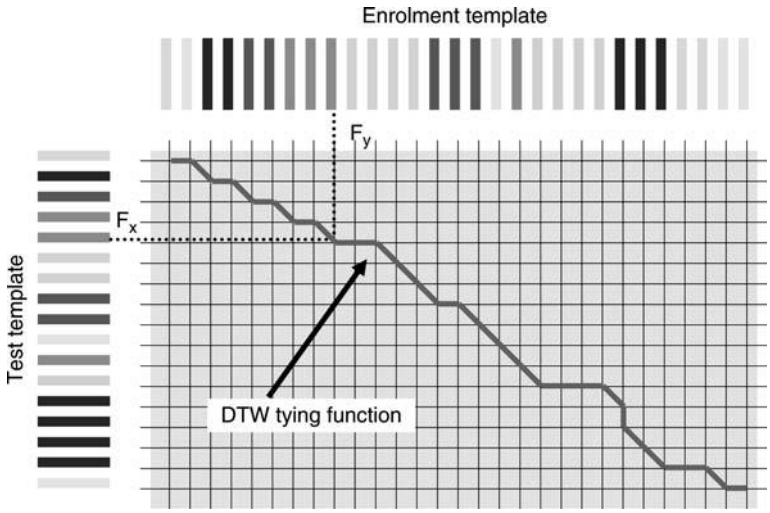
The speaker matching could be split into three main functionalities:

- Create a template from the feature set extracted from the enrollment data. Usually, the template is denoted "speaker model."
- Compare a feature set acquired from a sound captor with a speaker model and output a likelihood score.
- Take an identification decision using this score. Usually other information are used during this decision phase, like a model of potential impostors. In a speaker recognition system, the scores are very often normalized before to take the decision, using a ► Score Normalization technique.

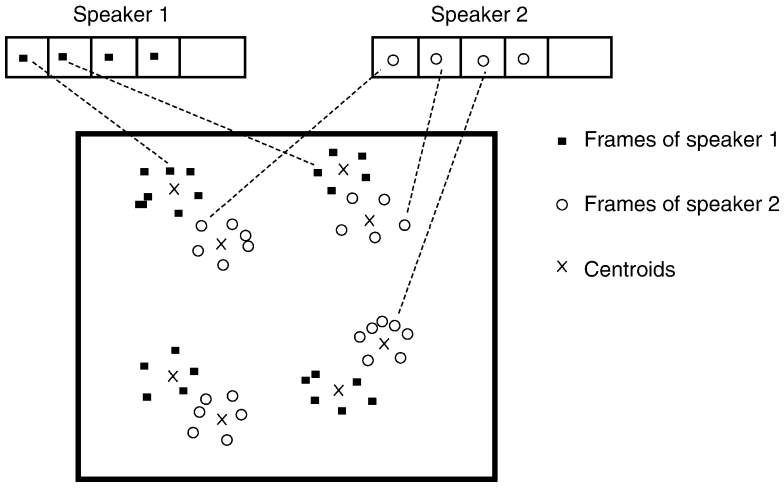
Introduction

Two main classes of approaches are usually used for speaker matching:

- Direct matching. This kind of approaches don't really use a modeling of the speaker voice: the enrollment voice sample is used directly as a model and a similarity function between two voice samples is used for the score computation. If the time synchronization aspects are taken into account, a dynamic time warping algorithm is used in order to find the best time alignment between the enrollment acoustic feature sequence and the test feature test sequence. ► Dynamic Time Warping (DTW) [1, 2] involves a strong dependency on the text pronounced by the individuals: the text should be the same during the enrollment and the recognition phases, the message should have a short duration (few seconds). The Fig. 1 illustrates the DTW main principles. For text independent systems, where the text pronounced by the individuals could be different between the enrollment and the recognition phases, a ► Vector Quantification (VQ) algorithm is used [2]. The main principle of VQ is illustrated in Fig. 2. In this case, the VQ codebook could be seen as a model, but it is in fact close to a data reduction of the enrollment sample: only a subset of the feature vectors extracted from the enrollment data are kept and the other are withdrawn. The main advantage of these methods is the low level of needed computer resources. These methods are also efficient in terms



Speaker Matching. **Figure 1** DTW principle.



Speaker Matching. **Figure 2** Frames distribution of two speakers in the acoustic space. Each speaker can be characterized by its own distribution modes (centroids).

of identification performance if the variability factors like utterance text content, microphone or environmental noises are controlled [1, 3].

- Machine learning approaches. In this class of methods, a speaker voice model is learnt using one or several enrollment recordings. During the test, the likelihood of the test data is computed using this model. Two kind of methods are used: the statistical modeling techniques (mainly GMM or HMM) and the discriminative classification techniques (mainly neural networks and SVM). Recently, mixed approaches were proposed, where a statistical approach (based on GMM) is used to deal with the data variability and a discriminative classifier is used for the decision estimation (SVM).

This chapter will describe more precisely the Machine Learning based approaches, which associate the GMM and the SVM.

GMM-UBM (GMM-MAP) Approach

GMM-UBM is the predominate approach used in speaker recognition systems, particularly for text-independent task [4]. This approach is based on a

generative statistical framework and follows the ▶ Bayesian Hypothesis Test representation.

This hypothesis test involves the estimation of two probabilities: ($H0$), Y comes from the hypothesized speaker S and ($H1$), Y is not from the hypothesized speaker S , where Y is the observed speech segment and S the targeted speaker. In GMM-UBM approach, the models are Gaussian Mixture Models which estimate a probability density function by:

$$p(x|\lambda) = \Sigma_{i=1}^{M} w_i N(x|\mu_i, \Sigma_i) \quad (1)$$

where w_i, μ_i and Σ_i are weights, means and covariances associated with the Gaussian components in the mixture. Usually a large number of components in the mixture and diagonal covariance matrices are used.

The model $\lambda_{\overline{hyp}}$ is denoted world model or Universal Background Model (UBM) when the model is environment independent. Its parameters are estimated using the EM algorithm, maximizing the Maximum Likelihood criterion. The speaker model λ_{hyp} parameters are generally obtained by adapting the world model parameters, using the Bayesian adaptation framework. Generally, only mean parameters are adapted and the other parameters remain unchanged [5]. The MAP adaptation procedure follows the formula:

$$\mu^i_{map} = \frac{n_i}{n_i + r} \cdot \mu^i_{emp} + (1 - \frac{n_i}{n_i + r}) \cdot \mu^i_{ubm} \quad (2)$$

where μ^i_{map} is the adapted mean for a given Gaussian component i, μ^i_{emp} is the corresponding empirical mean (obtained using the speaker enrollment data and EM algorithm), μ^i_{ubm} is the corresponding UBM mean, n_i is the occupancy value for the component (obtained also thanks to the EM algorithm, using the UBM and the enrollment data) and r is a regulation factor.

Speaker detection test relies on a log-likelihood ratio computation. Regarding the large size of the GMM models usually used (between 512 and 2,048 Gaussian components), a fast scoring technique is usually used. This technique consists in computing the ratio only on the n winning components, i.e. for a given frame only the n highest component likelihoods are computed for each target, the UBM model is used to find this top-component set. If W is the world model, f_s the test segment and L the hypothetic target model, the test is computed as follows (usually, $n = 10$):

$$l(f_s|L) = \sum_{i=1}^{n} l(f_s|\theta_i^L) + \sum_{i=n+1}^{G} l(f_s|\theta_i^W) \quad (3)$$

where L and W are GMMs of G gaussian components, each one of them respectively described by θ_i^L and θ_i^W. (θ being the parameters of a gaussian in a mixture: $\theta_i = (\mu_i, \sigma_i, \alpha_i)$ with $i \in [1, G]$). It is important to emphasize the multiple roles of the background model, the UBM. This model is used to represent the acoustic/phonetic/linguistic space. It is derived in order to obtain the speaker models (and only the mean parameters are adapted for these speaker models). The UBM model also drives the component selection during the testing phase (fast computation technique). Finally, in the decision step, the UBM represents the inverse hypothesis ($H1$: Y is not from the hypothesized speaker S). The UBM is clearly one of the key part of a GMM-UBM speaker recognition system.

GMM Supervector Linear Kernel (GSL)

The SVM approach offers an alternative classification strategy to the widely used GMM and has been investigated by many in the context of ASV, see for example [6, 7].

Recalling that ASV is a two-class problem then all expansion vectors corresponding to a given speaker in the training mode are labeled for example +1 and are confronted individually by expansions from a cohort of other speakers (loosely termed the impostor cohort) with the label −1. The result of the training is the definition of a separating hyperplane:

$$f(\mathbf{x}) = \sum_{i=1}^{N_{SV}} \alpha_i t_i \mathbf{R}^{-1/2} \Phi(X_i)\mathbf{x} + d \quad (4)$$

based on N_{SV} support vectors and where t_i represent the ideal output, $\sum_{i=1}^{N_{SV}} \alpha_i t_i = 0$, d is an offset and $\mathbf{R}^{-1}$ is a diagonal normalization matrix. Then the classifier model can be compacted as

$$\mathbf{w}_X = \begin{bmatrix} \sum_{i=1}^{N_{SV}} \alpha_i t_i \mathbf{R}^{-1} \Phi(X_i) \\ d \end{bmatrix} \quad (5)$$

enabling the evaluation of $f(\mathbf{x})$ with a simple dot product. Indeed, in the testing phase, the expansion of the test segment is augmented by the value 1 and

then a dot product between the two vectors of dimension $E+1$ is performed to produce a verification score:

$$Score_{SVM}(X, Y) = f(\mathbf{R}^{-1/2}\Phi(Y)) = [\Phi(Y)^t \ 1]\mathbf{w}_X. \quad (6)$$

Because $\mathbf{R}^{-1/2}$ is already integrated in $\mathbf{w}_X$, it is not required in the calculation of $f(\mathbf{R}^{-1/2}\Phi(Y))$.

The main difficulty for SVM based speaker-recognition is to obtain a fixed length input vector from a length-variable sequence of features. Several solutions were investigated, like the GLDS method proposed in [7]. Using the development of metrics in GMM space [8, 9] proposed to use the UBM-GMM system in order to extract the SVM input data. This solution combines the best of the two approaches: it takes advantage of the statistical modeling power of the GMM/generative approach and of the discriminative abilities of the SVM, which works only at the decision level. this approach is denoted GMM Supervector Linear Kernel (GSL) in this chapter. The SVM input vectors are gathered from the UBM-GMM parameters as defined bellow:

$$\Phi_{GSL}(X) = \mathbf{m}_X = \begin{bmatrix} \mathbf{m}_X^1 \\ \dots \\ \mathbf{m}_X^i \\ \dots \\ \mathbf{m}_X^C \end{bmatrix}, \quad (7)$$

It corresponds to the supervector comprising the values of means, $\mathbf{m}_X^i$, taken from the GMMs, trained on utterance X. Each GMM has C components and, with an acoustic feature vector of size F, this gives a $\Phi_{GSL}(X)$ of size CF. The weight and variance parameters from the UBM are used to define $\mathbf{r}$ with

$$\mathbf{r}_{GSL}^{-\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1}\Sigma_1^{-\frac{1}{2}} \\ \dots \\ \sqrt{\lambda_i}\Sigma_i^{-\frac{1}{2}} \\ \dots \\ \sqrt{\lambda_C}\Sigma_C^{-\frac{1}{2}} \end{bmatrix} \quad (8)$$

In terms of performance, the supervector approach like GSL is close to the GMM-UBM approach when a session mismatch technique is applied. Moreover, it allows to exploit other sources of information, like the information gathered from the GMM weights in [10].

Hidden Markov Model (HMM)

A Hidden Markov Model (HMM) is a double stochastic process in that it has an underlying stochastic process that is not observable (hence the term hidden) but can be observed through another stochastic process that produces a sequence of observations [11]. A Markov chain consists of states and arcs between these states. The arcs, which are associated to transition probabilities, permit to pass from one state to the another, to skip a state, or contrary to remain in a state. In a Hidden Markov Model, the real states sequence is hidden but the state sequence that minimize the probability of the observations given the HMM parameters could be easily determined using external observations, such as the vectors resulting from the pre-processing phase. The Hmms are more often used in text-dependent speaker recognition tasks, where there is a prior knowledge of the textual content. The Hmms have a theoretical advantage on the GMM, as Hmms can better model temporal variations [12, 13]. They also control the linguistic nature of the test speech segment, adding a kind of password-based security to the voice biometric identity verification. HMM-based methods have been shown to outperform conventional methods in text-dependent speaker verification [14].

Session Effects on Speaker Matching

The mismatch between the enrolment speech recording and the test speech recording is one of the main problem adressed in speaker recognition. Several factors compose this mismatch: the recording environment (room acoustic, other people, cars, TV, etc.), the microphone, the signal transmission channel, the phonetic or linguistic content of the messages, the pathological aspects of the speaker or the voice aging (for a given person, the voice is changing along the life).

Dealing correctly with the session mismatch problem is mandatory in order to obtain a robust speaker recognition system. The chapter *Session effects on speaker modelling* is dedicated to this subject.

Related Entries

- ▶ Gaussian Mixture Model
- ▶ Hidden Markov model
- ▶ Speaker Features
- ▶ Universal Background Model

References

1. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE T. Acoust. Speech (ASSP-26) **26**(1), 43–49 (1978)
2. Booth, I., Barlow, M., Watson, B.: Enhancements to dtw and vq decision algorithms for speaker recognition **13**(3–4), 427–433 (1993)
3. Soong, F.K., Rosenberg, A.E., Rabiner, L.R., Juang, B.H.: A vector quantization approach to speaker recognition. Approach Speaker Recogn., **66**(2), 14–26 (1987)
4. Bimbot, F., Bonastre, J.F., Fredouille, C., Gravier, G., Magrin-Chagnolleau, I., Meignier, S., Merlin, T., Ortega-Garcia, J., Petrovska, D., Reynolds, D.A.: A tutorial on text-independent speaker verification. EURASIP Journal on Applied Signal Processing, Special issue on biometric signal processing (2004)
5. Reynolds, D.A., Quatieri, T.F., Dunn, R.B.: Speaker verification using adapted Gaussian mixture models. Digit. Signal Process, **10**(1–3), 19–41 (2000)
6. Wan, V.: Speaker Verification Using Support Vector Machines. Ph.D. thesis, University of Sheffield (2003)
7. Campbell, W., Campbell, J., Reynolds, D., Singer, E., Torres-Carrasquillo, P.: Support vector machines for speaker and language recognition. Comput. Speech Lang., **20**(2–3), 210–229 (2006)
8. Ben, M., Betser, M., Bimbot, F., Gravier, G.: Speaker diarization using bottom-up clustering based on a parameter-derived distance between adapted gmms. In: ICSLP (2004)
9. Campbell, W.M., Sturim, D., Reynolds, D.A.: Support vector machines using GMM supervectors for speaker verification. IEEE Signal Process. Lett. **13** (2006)
10. Scheffer, N., Bonastre, J.F.: A UBM-GMM driven discriminative approach for speaker verification. In: Odyssey (2006)
11. Rabiner, L., Juang, B.: Fundamentals of Speech Recognition. Prentice-Hall, Upper Saddle River, (1992)
12. Nordstrm, T., Melin, H., Lindberg, J.: comparative study of speaker verification systems using the polycost database. In: International Conference on Spoken Language Processing ICSLP (1992)
13. Tishby, N.: On the application of mixture ar hidden markov models to text-independent speaker recognition. pp. 563–570 (1991)
14. Reynolds, D., Carlson. B.: Text-dependent speaker verification using decoupled and integrated speaker and speech recognizers. In: EUROSPEECH in Madrid, ESCA (1995)

Speaker Model

Speaker model is a representation of the identity of a speaker obtained from a speech utterance of known origin. It can be generative or discriminative. Most popular generative speaker models are the Gaussian Mixture Models (GMM), which model the statistical distribution of speaker features with a mixture of Gaussians. Typical discriminative speaker models are based on the use of Support Vector Machines (SVM), where the speaker model is basically a separating hyperplane in a high-dimensional space. Once enrolled, speaker models may be compared to a set of features coming from an utterance of unknown origin, to give a similarity score.

▶ Speaker Features

Speaker Parameters

▶ Speaker Features

Speaker Recognition Engine

▶ Speaker Matching

Speaker Recognition, One to One

▶ Liveness Assurance in Voice Authentication

Speaker Recognition, Overview

Jean Hennebert
Department of Informatics, University of Fribourg, Fribourg, Switzerland
Institute of Business Information Systems HES-SO Valais, TechnoArk, Sierre, Switzerland

Synonyms

Voice recognition; Voice biometric

Definition

Speaker recognition is the task of recognizing people from their voices. Speaker recognition is based on the extraction and modeling of acoustic features of speech that can differentiate individuals. These features conveys two kinds of biometric information: physiological properties (anatomical configuration of the vocal apparatus) and behavioral traits (speaking style). Automatic speaker recognition technology declines into four major tasks, *speaker identification*, *speaker verification*, *speaker segmentation*, and *speaker tracking*. While these tasks are quite different for their potential applications, the underlying technologies are yet closely related.

Introduction

Speaking is the most natural mean of communication between humans. Driven by a great deal of potential applications in human-machine interaction, automated systems have been developed to automatically extract the different pieces of information conveyed in the speech signal (Fig. 1). Speech recognition systems attempt to transcribe the content of what is spoken. Language identification systems aim at discovering the language in use. Speaker recognition systems aim to discover information about the identity of the speaker.

Interestingly, speaker recognition is one of the few biometric approach which is not based on image processing. Speaker dependent features are actually indirectly measured from the speech signal which is 1-dimensional and temporal. Speaker recognition is a biometrics qualified as *performance-based* or *active* since the user has to cooperate to produce a sequence of sounds. This is also a major difference with other *passive* biometrics such as for fingerprints, iris, or face recognition systems where user cooperation is not requested.

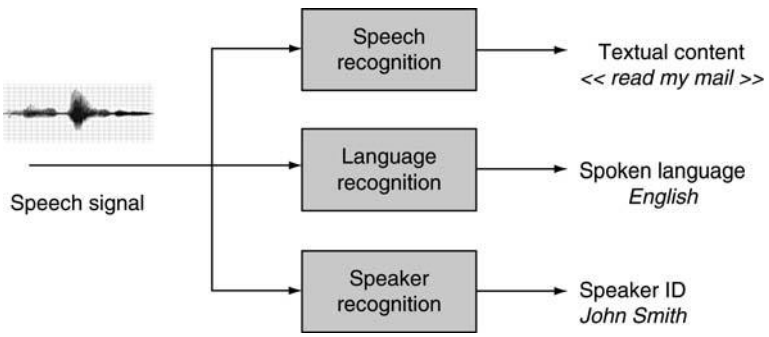
Speaker recognition technologies are often ranked as less accurate than other biometric technologies such as fingerprint or iris scan. However, there are two main factors that make voice a compelling biometric. First, there is a proliferation of automated telephony services for which speaker recognition can be directly applied. Telephone handsets are indeed available basically everywhere and provide the required sensors for the speech signal. Second, talking is a very natural gesture and it is often considered as lowly intrusive by users as no physical contact is requested. These two factors, added to the recent scientific progresses, made speaker recognition converge into a mature technology.

Speaker recognition finds applications in many different areas such as access control, transaction authentication, forensics, speech data management, and personalization. Commercial products offering voice biometric are available from different vendors. However, many technical and non-technical issues, discussed in the next sections, still remain open and are still subjects of intense research.

History of Speaker Recognition

Research and development on speaker recognition methods and techniques have now spanned more than five decades and it continues to be an active area [1].

In 1941, the laboratories of Bell Telephone in New Jersey produced a machine able to visualize spectrograph of voice signals. During the Second World War, the work on the spectrograph was classified as a military project. Acoustic scientists used it to attempt to



Speaker Recognition, Overview. **Figure 1** The different speech tasks can be declined into speech recognition, language identification, and speaker recognition.

identify enemy voices from intercepted telephone and radio communications. In the 1950's and 1960's, so-called *Experts* testimony in forensic application started. These experts were claiming that spectrographs were a precise way to identify individuals, which is of course not true in most conditions. They associated the term "voiceprint" to spectrographs, as a direct analogy to fingerprint [2]. This *expert* ability to identify people on the basis of spectrographs was very much disputed in the field of forensic applications, for many years and even until now [3].

The introduction of the first computers and mini-computers in the 1960's and 1970's triggered the beginning of more thorough and applied research in speaker recognition [4]. More realistic access control applications were studied incorporating real-life constraints as the need to build systems with single-session enrolment. In the 1980's, speaker verification began to be applied in the telecom area. Other application issues were then uncovered, such as unwanted variabilities due to microphone and channel. More complex statistical modelling techniques were also introduced such as the Hidden Markov Models [5]. In the 1990's, common speaker verification databases were made available through the Linguistic Data Consortium (LDC). This was a major step that triggered more intensive collaborative research and common assessment. The National Institute of Standards and Technology (NIST) started to organize open evaluations of speaker verification systems in 1997.

In the present decade, the recent advances in computer performances and the proliferation of automated system to access information and services pulled speaker recognition systems out of the laboratories into robust commercialized products. Currently, the technology remains expensive and deployment still needs lots of customization according to the context of use. From a research point of view, new trends are also appearing. For example, the extraction of higher-level information such as word usage or pronunciation is studied more for applications and new systems are attempting to combine speaker verification with other modalities such as face [6, 7] or handwriting [8].

Speech Signal

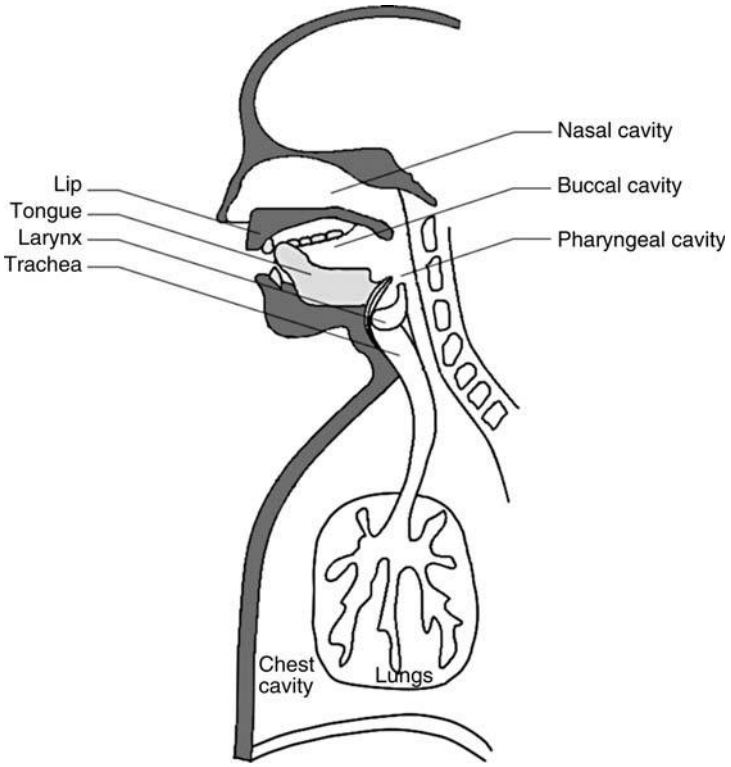
Speech production is the result of the execution of neuromuscular commands that expel air from the lungs, causes vocal cords to vibrate, or to stay steady and shape the tract through which the air is flowing out. As illustrated in Fig. 2, the *vocal apparatus* includes three cavities. The pharyngeal and buccal cavities form the *vocal tract*. The nasal cavity form the *nasal tract* that can be coupled to the vocal tract by a trap-door mechanism at the back of the mouth cavity. The vocal tract can be shaped in many different ways determined by the positions of the lips, tongue, jaw, and soft palate.

The *vocal cords* are located in the larynx and, when tensed, have the capacity to periodically open or close the larynx to produce the so-called *voiced sounds*. The air is hashed and pulsed in the vocal apparatus at a given frequency called the *pitch*. The sound then produced resonates according to the shapes of the different cavities. When the vocal cords are not vibrating, the air can freely pass through the larynx and two types of sounds are then possible: *unvoiced sounds* are produced when the air becomes turbulent at a point of constriction and *transient plosive sounds* are produced when the pressure is accumulated and abruptly released at a point of total closure in the vocal tract.

Roughly, the speech signal is a sequence of sounds that are produced by the different articulators changing positions over time [9]. The speech signal can then be characterized by a time-varying frequency content. Figure 3 shows an example of a voice sample. The signal is said to be slowly time varying or quasi-stationary because when examined over short time windows (Fig. 3-b), its characteristics are fairly stationary ($5-100$ msec) while over long periods (Fig. 3-a), the signal is non-stationary (>200 msec), reflecting the different speech sounds being spoken.

The speech signal conveys two kinds of information about the speaker's identity:

1. *Physiological properties*. The anatomical configuration of the vocal apparatus impacts on the production of the speech signal. Typically, dimensions of the nasal, oral, and pharyngeal cavities and the length of vocal cords influence the way phonemes are produced. From an analysis of the speech signal, Speaker recognition systems will indirectly capture some of these physiological properties characterizing the speaker.
2. *Behavioral traits*. Due to their personality type and parental influence, speakers produce speech with different phonemes rate, prosody, and coarticulation



Speaker Recognition, Overview. Figure 2 Schematic view of the human vocal apparatus. The vocal apparatus includes three cavities: the pharyngeal, buccal, and nasal cavities. These cavities form the vocal and nasal tract that can be shaped in many different ways determined by the positions of the lips, tongue, jaw, and soft palate.

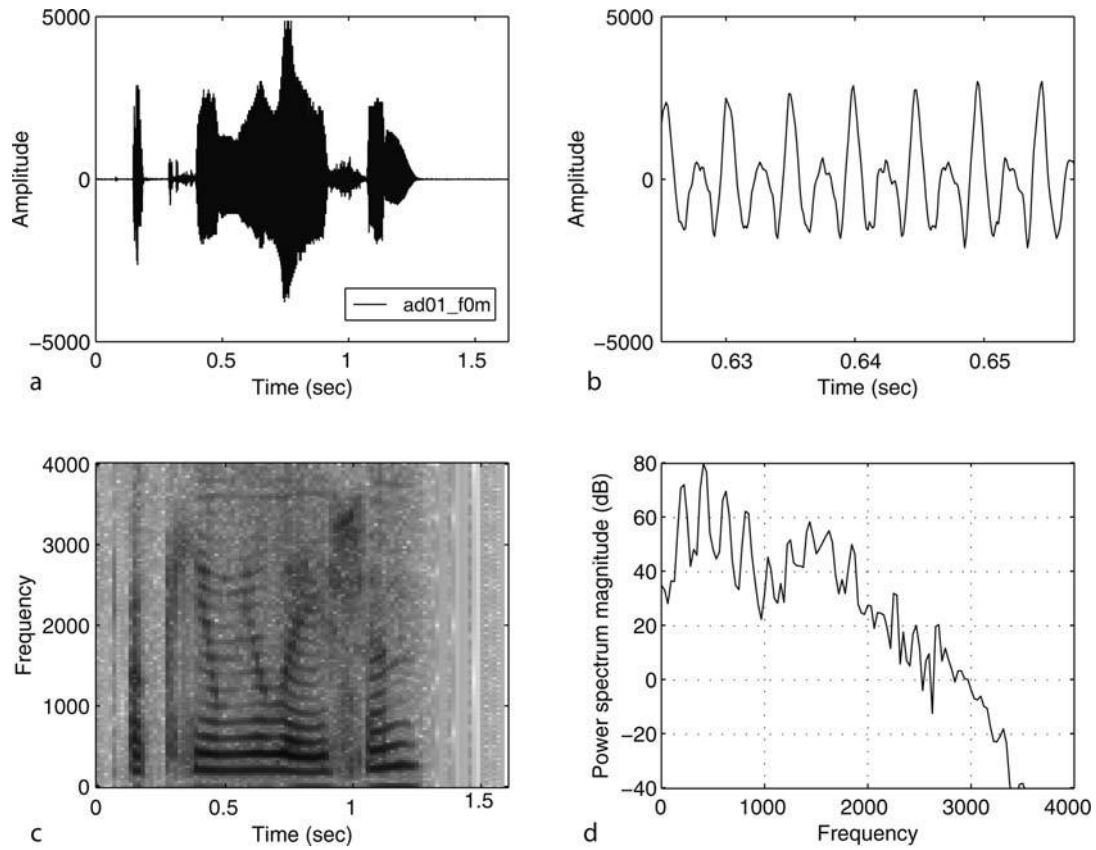
effects. Due to their education, socio-economic status, and environment background, speakers use different vocabulary, grammatical constructions, and diction. All these higher-level traits are of course specific to the speaker. Hesitation, filler sounds, and idiosyncrasies also give perceptual cues for speaker recognition.

Most of the speaker recognition systems are relying on low-level acoustic features that are linked to the physiological properties. Some behavioral traits such as prosody or phoneme duration are partly captured by some systems. Higher-level behavioral traits such as preferred vocabulary are usually not implicitly modeled by speaker recognition systems because they are difficult to extract and model. Typically, the system would need a large amount of enrolment data to determine the preferred vocabulary of a speaker, which is not reasonable for most of the commercial applications.

Intra-speaker variabilities are due to differences of the state of the speaker (emotional, health, . . .). Inter-speaker variabilities are due to physiological or behavioral differences between speakers. Automatic speaker recognition systems exploit inter-speaker variabilities to distinguish between speakers but are impaired by the intra-speaker variabilities which are, for the voice modality, numerous.

Feature Extraction and Modeling

In the case of the speech signal, the feature extractor will first have to deal with the long-term non-stationarity. For this reason, the speech signal is usually cut into frames of about 10-30 msec and feature extraction is performed on each piece of the waveform. Secondly, the feature extraction algorithm has to cope with the short-term redundancy so that a reduced and relevant acoustic information is extracted. For this purpose, the representation of the waveform is generally swapped from the temporal domain to the frequency domain, in which the short-term temporal periodicity is represented by higher energy values at the frequency



Speaker Recognition, Overview. **Figure 3** Speech signal of the word *accumulation*: (**a**) waveform, (**b**) partial waveform, (**c**) narrow-band spectrogram of (**a**), (**d**) power spectrum magnitude of (**b**).

corresponding to the period. Thirdly, feature extraction should smooth out possible degradations incurred by the signal when transmitted on the communication channel. For example, in the case of telephone speech, the limited bandwidth and the channel variability will need some special treatment. Finally, feature extraction should map the speech representation into a form which is compatible with the statistical classification tools in the remainder of the processing chain.

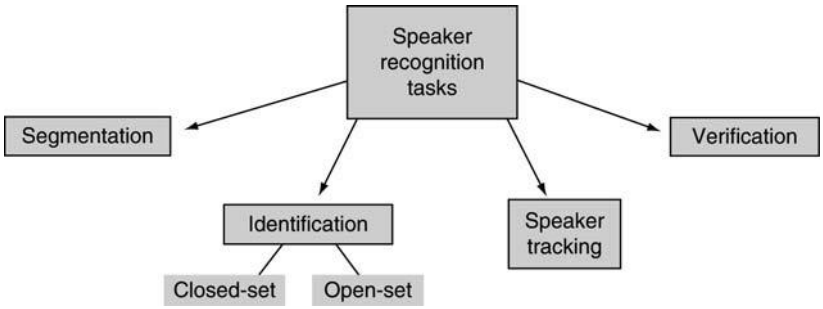
Usual feature extraction techniques are the so-called *linear predictive coding (LPC) cepstral* analysis or the *mel-frequency cepstral* analysis. These algorithms are widely used in the field of speech processing [9, 10]. The output of the feature extraction module is a temporal sequence of acoustic vectors $X = \{x_1, x_2, \ldots, x_N\}$ of length N with each vector x_n having a constant dimension D. The sequence X is then input into the pattern classification module.

There are many different ways reported in the scientific literature to build speaker models: vector quantization, second order statistical methods, Gaussian Mixtures Model (GMM), Artificial Neural Network (ANN), Hidden Markov Model (HMM), Support Vector Machines (SVM), etc. One of the most widely used is GMM modeling. By nature, GMMs are versatile as they can approximate any probability density function given a sufficient number of mixtures. With GMMs, the probability density function $p(x_n|M_{client})$ or *likelihood* of a D-dimensional feature vector x_n given the model of the client M_{client}, is estimated as a weighted sum of multivariate gaussian densities (e.g., [11]).

Speaker Recognition Tasks and Applications

Automatic speaker recognition can be declined into four tasks (Fig. 4).

Speaker identification attempts to answer the question "Whose voice is this?" In the case of large speaker



Speaker Recognition, Overview. Figure 4 From left to right, the different speaker recognition tasks can be loosely classified from the most difficult to the less difficult ones. The tasks of verification and identification are the major ones considering the potential commercial applications.

sets, it can be a difficult task where chances are more to find speakers with similar voice characteristics. The identification task is said to be *closed-set* if it is sure that the unknown voice comes from the set of enrolled speaker. By adding a "none-of-the-speaker" option, the task becomes an *open-set* identification. Speaker identification is mainly applied in surveillance domains and, apart from this, it has a rather small number of commercial applications. *Speaker verification* (Also known as *speaker detection* or *speaker authentication* task.) attempts to answer the question "Is this the voice of Mr Smith?" In other words, a candidate speaker claims an identity and the system must accept or reject this claim. Speaker verification has a lot of potential commercial applications thanks to the growing number of automated telephony services. When multiple speakers are involved, these tasks can be extended to *speaker tracking* (when a given user is speaking) and *speaker segmentation* (blind clustering of a multi-speaker record).

Speaker recognition systems can also be classified according to the type of text that the user utters to get authenticated. One can distinguish between ► *text-dependent*, ► *text-prompted*, and ► *text-independent* systems. These categories are generally used to classify speaker verification tasks. To some extent, they can also apply to the task of identification.

- *Text-dependent systems.* These systems use the same piece of text for the enrolment and for the subsequent authentication sessions. Recognition performances of text-dependent systems are usually good. Indeed, as the same sequence of sounds is produced from session to session, the characteristics extracted from the speech signal are more stable. Text-dependency also allows to use finer modeling techniques capable to capture information about sequence of sounds. A major drawback of text-dependent systems lies in the replay attacks that can be performed easily with a simple device playing back a pre-recorded voice sample of the user. The term *password-based* is used to qualify text-dependent systems where the piece of text is kept short and is not supposed to be known to other users. There are *system selected text/password* where an a priori fixed phrase is composed by the system and associated to the user (e.g., pin codes) and *user selected text/password* where the user can freely decide on the content of the text.
- *Text-prompted systems.* Here the sequence of words that need to be said is not known in advance by the user. Instead, the system prompts the user to utter a randomly chosen sequence of words. A text-prompted system actually works in two steps. First, the system performs speech recognition to check that the user has actually said the expected sequence of words. If the speech recognition succeeds, then the verification takes place. This *challenge-response* strategy achieves a good level of security by preventing replay attacks.
- *Text-independent.* In this case, there is no constraint on the text spoken by the user. The advantages are the same as for the text-prompted approach: no password needs to be remembered and the system can incrementally ask for more data to reach a given level of confidence. The main drawback lies here in the vulnerability against replay attacks since any recording of the user's voice can be used to break into the system.

Speaker recognition finds applications in many different areas such as telephony transaction authentication, access control, speech data management, and forensics. It is in the telephony services that speaker recognition finds the largest deal of applications as the technology can be directly applied without the need to install any sensors.

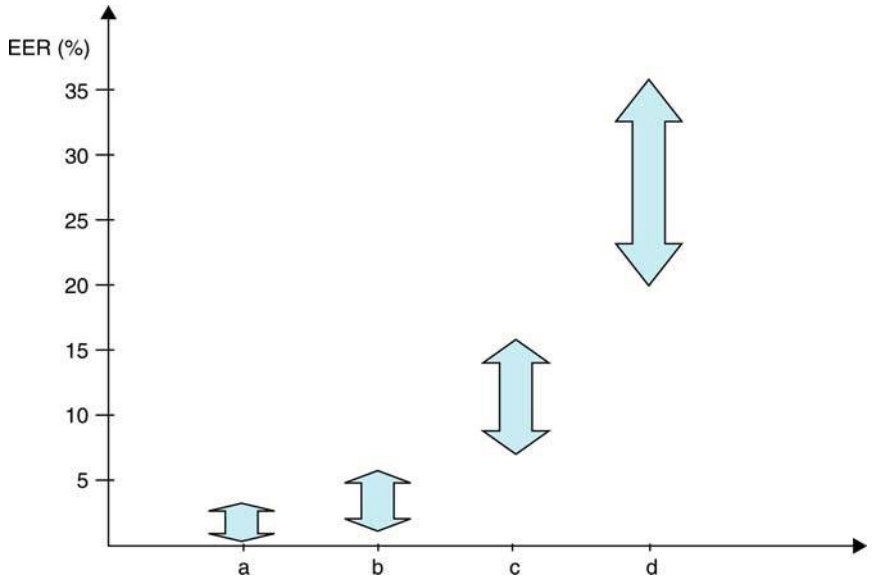
- *Telephony authentication for transactions.* Speaker recognition is the only biometric that can be directly applied to the automated telephony services (Interactive Voice Response - IVR systems). Speaker recognition technology can be used to secure the access to reserved telephony services or to authenticate the user while doing automated transactions. Banks and telecommunication companies are the main potential clients for such systems. As many factors impact on the performances of speaker recognition in telephony environment, it is often used as a complement to other existing authentication procedures. Most of the implementations are using a text-prompted procedure to avoid pre-recording attacks and to facilitate the interaction with a dialog where the user just needs to repeat what the system is prompting. A less known but interesting example of speaker verification application in telephony is also the home incarceration and parole/probation monitoring.
- *Access control.* Speaker verification can be used for physical access control in combination with the usual mechanisms (key or badge) to improve security at relatively low cost. Applications such as voice-actuated door locks for home or ignition switch for automobile are already commercialized. Authorized activation of computers, mobile phones, or PDA is also an area for potential applications. Such applications are often based on text-dependent procedures using single passwords.
- *Speech data management and personalization.* Speaker tracking can be used to organize the information in audio documents by answering the questions: who and when a given speaker has been talking? Typical target applications are in the movie and media industry with speaker indexing and automatic speaker change detection for automatic subtitling. Automatic annotation of meeting recordings and intelligent voice mail could also benefit from this technology. In the area of personalization, applications to recognize broad speaker characteristics such as gender or age can be used to personalize advertisements or services.
- *Forensic speaker recognition.* Some criminal cases have recordings of lawbreakers voice and, speaker verification technologies can help the investigator in directing the investigation. On the other hand, there is a general acceptation in the scientific community on the fact that a verification match obtained with an automatic system or even with a so-called voiceprint expert, should not be used as a proof of guilt or innocence [3].

Performances and Influencing Factors

Figure 5 summarizes typical ranges of Equal Error Rate (EER) performances for four categories of speaker verification systems [12]. The range of performances is globally extremely large, going from 0.1 to 30% across the systems. Text-dependent applications using high quality speech signals can have very low EER typically ranging from 0.1 to 2%. Such performances are obtained with multi-session enrolment of several minutes and test data of several seconds acquired in the same condition as for the enrolment. Pin-based text-dependent applications running on the telephony channel will typically show performances ranging from 2 to 5%. Text-independent applications based on telephony quality, recorded during conversations over multiple handsets and using several minutes of multi-session enrolment data and a dozen of seconds for the test data, will show EER ranging from 7 to 15%. Finally, text-independent applications based on very noisy radio data will show performances ranging from 20 to 35%.

Summary

Speaker recognition is often ranked as providing medium accuracy in comparison to other biometrics. This is due to three main factors. First, there are the inherent and numerous intra-speaker variabilities of the speech signal (emotional state, health condition, age). Second, the inter-speaker variabilities are



Speaker Recognition, Overview. Figure 5 Typical performances of speaker verification systems. The arrows define ranges of Equal Error Rates for four different types of applications. Applications of type (**a**) are text-dependent based on high quality speech signals. Applications of type (**b**) are text-dependent based on telephony speech quality, typically a pin-based application. Applications of type (**c**) are text-independent on telephony speech quality recorded during conversations. Applications of type (**d**) are text-independent based on very noisy radio.

relatively weak, especially within family members. Finally, the speech signal is often exposed to all sort of environmental noise and distortions due to the communication channel. These varying acquisition conditions are captured by the speech template which becomes biased. To smooth out these variabilities, lengthy or repeated enrollment sessions are often performed, but this is generally at the expense of usability.

Speaker recognition remains however a compelling biometrics. First, talking is considered a very natural gesture and user acceptance is generally high. Furthermore no physical contact is requested to record the biometric sample and the rate of failure to enroll is also very low. Finally, the technology cost of ownership is pretty low. For computer-based applications, simple sound cards and microphones are available at low-cost. For telephony applications, there is no need for special acquisition devices as any handset can be used from basically anywhere.

Speaker recognition technology has made tremendous progress over the past 20 years and finds new applications in many different areas such as telephony authentication, access control, law enforcement, speech data management, and personalization.

Related Entries

- ▶ Biometrics, Overview
- ▶ Speaker Feature
- ▶ Session Effects on Speaker Modeling
- ▶ Speech Analysis
- ▶ Speech Production

References

1. Furui, S.: 50 years of progress in speech and speaker recognition. In: Proceedings of SPECOM, pp. 1–9 (2005)
2. Kersta, L.: Voiceprint Identification. Nature **196**, 1253–1257 (1962)
3. Boe, L.J.: Forensic voice identification in France. Speech Commun. **31**, 205–224 (2000)
4. Atal, B.S.: Automatic recognition of speakers from their voices. Proc. IEEE **64**, 460–475 (1976)
5. Naik, J.M., Netsch, L.P., Doddington, G.R.: Speaker verification over long distance telephone lines. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Glasgow, Scotland pp. 524–527 (1989)
6. Jain, A., Ross, A., Prebhakar, S.: An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics **14**(1) (2004)

7. Fauve, B., Bredin, H., Karam, W., Verdet, F., Mayoue, A., Chollet, G., Hennebert, J., Lewis, R., Mason, J., Mokbel, C., Petrovska., D.: Some results from the biosecure talking face evaluation campaign. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Las Vegas, USA (2008)
8. Humm, A., Hennebert, J., Ingold, R.: Spoken signature for user authentication. SPIE J. Electron. Imaging, Special Section on Biometrics: ASUI **17**(1) (2008)
9. Rabiner, L., Juang, B.H.: Fundamentals of Speech Recognition. Prentice Hall (1993)
10. Picone, J.: Signal modeling techniques in speech recognition. Proc. IEEE **81**(9), 1214–1247 (1993)
11. Reynolds, D.: Automatic speaker recognition using gaussian mixture speaker models. Linc. Lab. J. **8**(2), 173–191 (1995)
12. Reynolds, D.: An overview of automatic speaker recognition technology. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, pp. 4072–4075 (2002)

Speaker Recognition, Standardization

Judith Markowitz
Consultants, Chicago, IL, USA

Synonyms

Speaker authentication; Speaker biometrics; Speaker identification and verification, SIV; Voice authentication; Voice recognition

Definition

The term "speaker recognition" (SR) refers to a group of technologies that use information extracted from a person's speech to perform biometric operations such as speaker identification and verification (SIV). Standards for SR are designed to support the development of applications that can work with technology from different vendors (application programming interface standards), the sharing of SR data (data interchange standards), the transmission of data in real time (distributed speaker recognition standards), and the management of data resources in distributed environments (process-control protocol standards).

Introduction

SR technologies stand at the juncture between ▶ speech-processing and biometrics. They belong in speech processing, because they extract and analyze data from the ▶ stream of speech. They belong in biometrics, because the data that are extracted describe a physical or behavioral characteristic of the speaker and because they use that information to make decisions regarding the speaker, usually determining the identity of the speaker and verifying a claim of identity. Some SR technologies perform other speaker-related functions, such as placing the speaker into a category, such as female or male (▶ speaker classification); determining whether the speaker has changed (speaker change); assessing the speaker's level of stress or emotion (emotion detection, voice stress analysis); tracking a specific voice in a multispeaker communication (speaker/voice tracking); separating interleaved and overlapping voices from each other (▶ speaker separation); and determining whether the speaker is lying or telling the truth (voice lie detection).

Standards for SR come from both speech processing and from biometrics. They fall into several categories:

1. Application programming interface (API) standards,
2. Sharing of stored SR data (data interchange),
3. Transmission of data in real time (distributed speaker recognition) and
4. Management of data resources in distributed environments (process-control protocols).

Application Programming Interface (API) Standards – Early Work

API standards eliminate the need for programmers to learn a new set of programming functions for each SR product. They accomplish this by establishing a standard set of functions that can be used to develop applications using any standards-compliant SR technology.

The bulk of the work on SR standards has been directed toward the development of standard APIs. Most of these standards have been crafted by speech-processing industry consortia or standards bodies and are extensions of existing standards for ▶ speech recognition.

The first and, to date, the most detailed API standard is the Speaker Verification API (SVAPI) [1, 2]. SVAPI was constructed by a speech- and biometrics-industry consortium formed in 1996, whose work was sponsored by Novell Corporation. The goal was to develop a companion to Speech Recognition API (SRAPI), an API standard for speech recognition on the PC desktop.

SVAPI is a low-to-midlevel standard that covers enrollment, verification, identification, and speaker classification with some support for speaker separation. It handles both centralized and distributed deployments and includes the functionality for specifying features of the stream of speech, the inclusion of several types of normalized scoring, and the characterization of input from both microphones and telephones. SVAPI consists of a set of callable Dynamically Linked Library (DLL) functions. It is written in C++ and Java and runs under Windows on desktop platforms.

SVAPI 1.0 was released in 1997, but work on the specification stopped shortly thereafter and the standard remains largely unsupported. Despite its short life as a standard, SVAPI has had a lasting impact on API standards in both biometrics and speech processing.

Work on SVAPI inspired the development of a high-level, generic API for biometrics that was developed by The National Registry, Inc. (NRI) under contract with an agency of the US Department of Defense. The resulting specification called Human Authentication API (HA-API) [3] was the precursor to the BioAPI specification of the BioAPI Consortium (www.bioapi.org). Proof-of-concept testing began early in 1998 and was performed on five commercial biometric products, including one SR product.

HA-API was designed for desktop platforms running 32-bit Windows operating systems. It supported stand alone and client–server implementations. HA-API operations required by SR, such as adaptive updating voice models (called "adaptation"), were retained when HA-API evolved into the BioAPI specification.

The *S.100 Media Resources and Service Protocol* [4] was developed by the Enterprise Computer Telephony Forum (ECTF). It was an API standard for using speech recognition (ASR) in computer-telephony. Support for speaker verification and identification was added to S.100 version 2 in the form of two parameters: ASR_ECTF_Verification and ASR_ECTF_Identification. Their role was to extend the functionality of speech recognition (ASR) technology. "When supported, the ASR resource may be used for speaker identification and speaker verification, e.g., by training a context with ► utterances from a particular speaker [4]."

API Standards – Current Work

VoiceXML is an XML scripting language for developing speech applications for ► interactive voice response (IVR) applications over the telephone. It is the dominant standard for ASR and text-to-speech synthesis. Its developers, the World Wide Web Consortium (W3C) and the VoiceXML Forum, are defining an SR module for the next version of VoiceXML (version 3.0). The Forum's Speaker Biometrics Committee (SBC) has identified the requirements for the module, and the W3C's Voice Browser Working Group is constructing the actual specification. When completed, the SR module will become part of a network of speech-processing standards for services-oriented architecture. Figure 1 shows the network and indicates where the SR module (called "SIV") will fit.

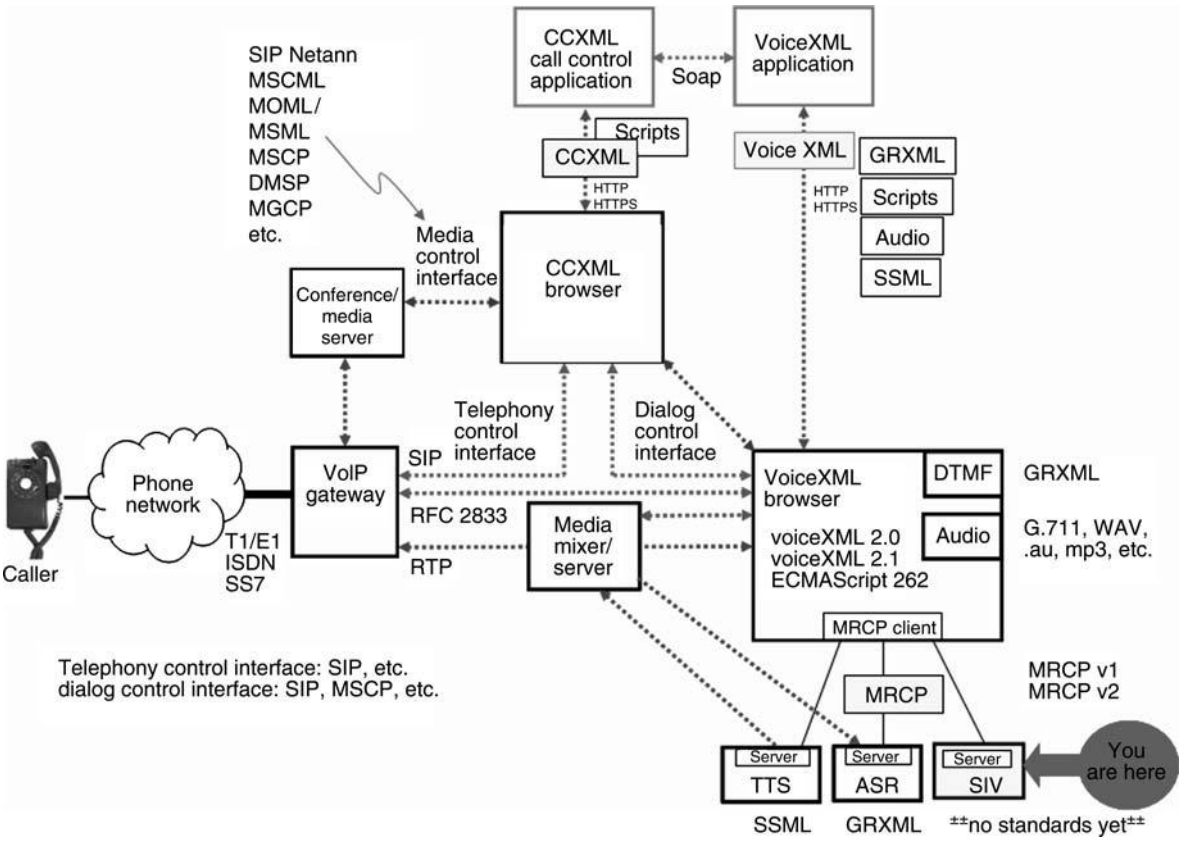
The SBC has published several documents related to its work on the Forum's web page (http://www.voicexml.org/biometrics):

1. SIV Glossary [6],
2. SIV Applications [7] a review of existing and potential SR applications, and
3. *Speaker Identification and Verification (SIV) Requirements for VoiceXML Applications* [8].

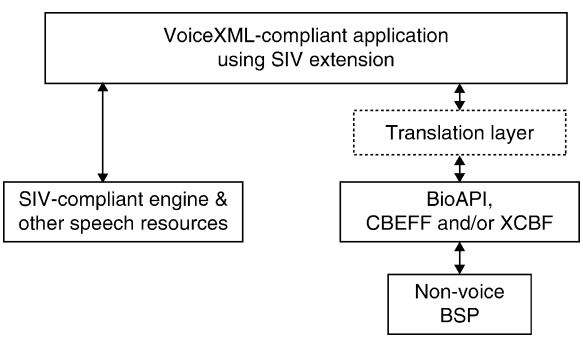
The requirements document specifies the basic functions that an API standard for speaker verification and identification must support: enrollment, verification, identification, and supervised adaptation. It also establishes a three-phase "session" as the basic unit of operation. Those phases are

1. Designation – when the function is specified (e.g., verification), and preparatory events occur (e.g., claim of identity)
2. Audio processing – when speech samples are collected and analyzed and decisions are rendered
3. Cleanup – when temporary files and data are purged and the session is concluded.

The document denotes a set of basic properties, including various kinds of thresholds (e.g., decision



Speaker Recognition, Standardization. Figure 1 Network of standards for speech processing [5].



Speaker Recognition, Standardization. Figure 2 Relationship of a VoiceXML module for SR and the BioAPI specification.

threshold, adaptation threshold), timeouts, and limits (e.g., minimum number of utterances required to perform verification). It defines allowable concurrent and nested sessions and provides support for multifactor applications. As shown in Fig. 2, the requirements document also indicates how a VoiceXML SR module might work with the generic biometric standard, the BioAPI specification.

Data Interchange Standards

Data exchange/interchange standards support the sharing and reuse of enrollment, verification, and identification data. They are needed by a broad spectrum of operations requiring interoperability, such as product upgrades that are not backwards compatible, security audits, inter-bank customer support, and multiagency intelligence and law-enforcement investigations. They facilitate the exchange of SR data by providing a structure that not only transmits the data but also offers a controlled description of those data. The data exchanged by a data interchange standard may be raw, partially-processed/feature data, or fully-processed model/template data.

The VoiceXML Forum and Technical Committee M1 (Biometrics) of the InterNational Committee for

Speaker Recognition, Standardization. **Table 1** Representative Elements of SIVR-1 Session Header

Name	Status	Data Type	Value(s)
Purpose	Required	String	Verification
			Identification
			Enrollment
			Multiple
			Other
Channel	Required	Complex type	
AudioFormatHeader	Required	Complex type	
Security	Optional	Complex type	
Speaker	Optional	Complex type	
Input device	Optional	Complex type	

Information Technology Standards (INCITS) [9] are collaborating on the development of an American National Standard for SR. As its name suggests, the draft standard *Speaker Recognition Format for Raw Data Interchange (SIVR-1)* supports the interchange of raw SR data. SIVR-1 is a format for describing the data being transmitted for a single SR session. It supports enrollment, verification, and identification operations. Work on SIVR began in 2005 and the standard is currently wending its way through separate approval processes by INCITS and the VoiceXML Forum.

SIVR-1 defines two headers: "Session" and "Instance." It also supports the inclusion of nonstandardized data (called "extended" data). Since SIVR-1 is an XML standard, it also specifies an XML schema.

Each SIVR-1 compliant format has a single Session header. Table 1 contains a subset of the XML elements included in the Session header. The Session header contains information that remains constant throughout the session. Those elements include the date and time the session took place, the total amount of utterance data included in the session, characteristics of the channel and input device, and a description of the data.

These elements are governed by existing standards. For example, the syntax of the DateAndTime element must comply with ISO 8601 2004(E) *Data Elements and Interchange Formats – Interchange Formats - Representation of Dates and Times* [10]. Although security is essential for protecting data stored in an SIVR-1 format, that element is optional to avoid conflict with external security and identity management technologies that may be applied.

Speaker Recognition, Standardization. **Table 2** Elements in audio format header

Name	Status	Data Type	Value(s)
Byte Order	Required	hexBinary	0Xff00
Streaming	Required	boolean	0 or 1
AudioFormat	Required	string	LinearPCM
			Mu-Law
			A-Law
			OGG Vorbis
			OGG Stream
Samplingrate	Required	Integer	Samples per second
BitsPer Sample	Required	Integer	

Elements in Table 1 that are "complex type" are themselves made up of elements. Table 2 displays several of the elements that make up the AudioFormatHeader, which defines the data that are stored in and transmitted using SIVR-1.

The element AudioFormat specifies the audio formats to be used to store data in the format. These audio formats are widely used open standards.

An SIVR-1 format must contain at least one Instance Header. Each Instance contains information that can change from one of the speaker's utterances to the next within the Session. Instances also contain the raw data of the utterance. Table 3 displays some of the elements in the SIVR-1 Instance header.

As Table 3 reveals, each Instance in a Session is assigned a number. SIVType is included, because

Speaker Recognition, Standardization. Table 3 Representative Elements of SIVR-1 Instance Header

Name	Status	Data Type	Value(s)
Instance number	Required	Integer	
SIVType	Required	String	Text-dependent
			Text –prompted
			Text-independent
			Unknown
ASRUsed	Required	String	Yes
			No
			Unknown
Type of Prompt Content	Required	String	None
			Text
			Binary
			Pointer
			Both
Utterance	Required	Complex type	

different Instances can utilize different kinds of SR technology as the following example illustrates.

Instance 1 Investigator: "Please say your rank."
Speaker: "Corporal"
Instance 2 Investigator: "Please say your ID number"
Speaker: "7398722"
Instance 3 Investigator: "Where were you on the night of March 5, 2007?"
Speaker: "I was home alone."

ASRUsed is included, because some of the instances may use ASR, while other instances may not. It is more likely, for example, that ASR would be used for instances 1 and 2, which are the ► text-prompted and ► text-dependent technology than for instance 3, which is text-independent and requires a different type of ASR.

As with the elements of the Session header, the complex type element Utterance consists of several other elements that include the quality of the raw audio data and audio-format information that deviates from the default values specified in the AudioFormatHeader element of the Session.

In 2007, the Joint Technical Committee 1 (JTC 1) of the International Standards Organization (ISO) and the International Electrotechnical Commission (IEC) approved a project for the development of an international standard under JTC1 Subcommittee 37 – Biometrics (No. 1.37.19794–13, *Voice data*) that is similar in scope to the INCITS/VoiceXML project. This project differs from the INCITS/VoiceXML project in that it is developing binary and XML versions and will have header for standardized feature as wel as far raw data.

Distributed Speaker Recognition Standards

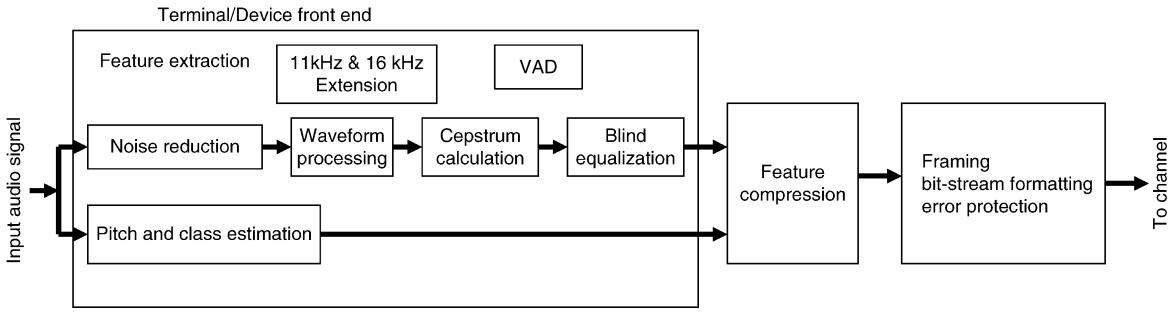
The real-world conditions under which SR must operate are not always optimal. SR data are captured by an increasingly diverse spectrum of heterogeneous, third-party input devices that process the data, using one of a growing number of standard audio formats so that they can be transmitted over telecommunications and data networks. Those networks (called "channels") differ in their acoustic characteristics, bandwidths, and quality. These variables affect the performance of even the most accurate SR technology and are represented in the data-interchange headers presented in the previous sections.

One method for reducing the impact of differences in data quality and processing associated with input devices and channels is to embed technology into input devices that perform standardized preprocessing and feature extraction before the data are sent over the channels. If those operations produce the features that are needed to perform speech recognition or SR, the embedded technology is called "distributed speech/speaker recognition" (DSR).

In 2000, the European Telecommunications Standards Institute (ETSI) published a standard for extracting those common features [11] in support of speech recognition. As Fig. 3 indicates, the embedded technology ("terminal front end") performs error-reduction, noise reduction, compression, and other operations in addition to feature extraction before transmitting the data.

Work is now being done to extend ETSI DSR to SIV. In order to accomplish that, several additional features need to be extracted from the speech signal [12].

Since most developers of speech recognition and SR technology use a core set of common features, the development of a DSR standard seems reasonable. The problem facing ETSI DSR and other DSR standards is



Speaker Recognition, Standardization. **Figure 3** ETSI Distributed Speech Recognition (DSR).

that each speech recognition and SR vendor approaches feature extracting in a unique way and those differences are considered to be part of the vendor's "secret sauce."

Process-Control Protocols

Process control/data transport standards facilitate real-time communications among the disparate elements of a system. They enable applications, servers, input devices, and SR technology to exchange data in real time quickly, effectively, and smoothly. This is particularly important in the burgeoning web-services/services-oriented architecture (SOA) environment, which often involves complex network interactions among different kinds of "nodes." Those nodes include devices (e.g., telephones), resources (e.g., an SR product), applications, and servers.

One standard that is used to support SR in SOA is the *Simple Object Access Protocol* (SOAP). SOAP is a W3C standard for exchanging messages (called a "data transport protocol") that can be used over HTTP and HTTPS (for secured transport). It allows one network node (e.g., a client) to send a message to another node (e.g., a browser or server) and to get an immediate response. SOAP is a generic data-transport protocol; it makes no mention of SR and does not address issues of special concern to transmission of audio data.

Unlike SOAP, the *Media Resources Control Protocol* (*MRCP*) was created specifically to control voice-related resources and to support the transport of speech data in SOA and IVR environments. As shown in Fig. 4 , MRCP mediates between the servers that house the speech and SR technologies (called "media processing resources") and applications or other entities on the network (called "clients") that need to communicate with them. Figure 4 provides a more detailed view of the architecture of MRCPv2.

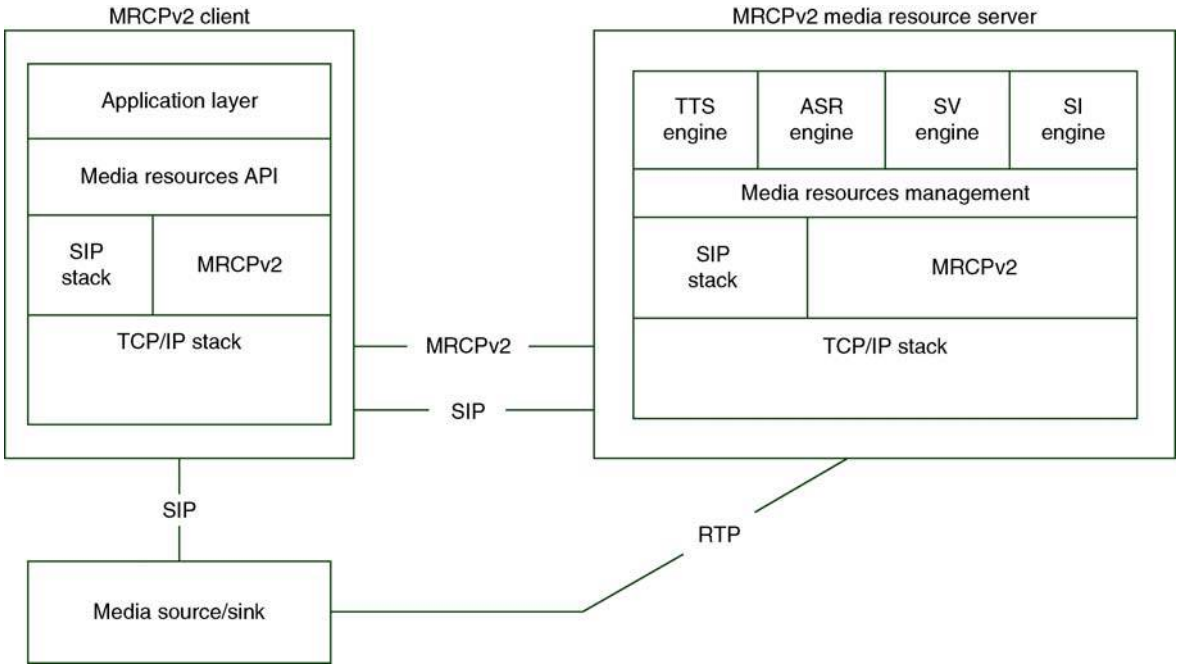
MRCPv2 specifies the messages that can be sent between the two parties, how the resources are to be used, and how these messages are to be carried over a transport layer. Figure 4 shows the two parties involved in the communication (client, resource server); the speech-processing resources that may be involved; and how the *Session Initiation Protocol* (SIP), the *Transmission Control Protocol* (TCP), and *Real-Time Transport Protocol* (RTP) are utilized.

An interaction between a client and a media resource server is called a "session."

A separate session may be created for each resource (e.g., a speaker-verification product and an ASR product) or a single session may involve multiple resources. For example, it supports the establishment of a single session for ASR and speaker verification that allows both resources to operate on the same utterances. The client uses Session Initiation Protocol (SIP) to start and end sessions and to establish an MRCP control channel with the media server so that the client can use the server's media processing resources. Once that is accomplished, MRCP-compliant messages can be sent between the client and the server. The SIP-labeled line between the client and the server, as shown in Fig. 4, indicates that SIP is also used to ensure that messages and audio are properly sent and received.

The commands/functions for speaker verification and identification are the "messages" that enable the client to control the SR operation within the session. They include commands to start and end sessions, to verify, identify, and get intermediate-level results.

MRCP is based on a requirements document that includes speaker verification and identification among the technologies to be supported [13], even though those technologies were not incorporated into MRCP



Speaker Recognition, Standardization. **Figure 4** MRCPv2 Architecture.

version 1. SR has been added to version 2, which also addresses security considerations, primarily for SR sessions.

MRCP version 1 (MRCPv1) was developed jointly by Cisco Systems, Inc., Nuance Communications, and Speechworks Inc. and has become a widely used standard within the speech-processing industry[14]. MRCPv2 was created by a speech-industry consortium within the Internet Engineering Technology Forum (IETF) and is in its final stages of approval [15].

Related Entries

► Biometrics, Overview
► Common Biometric Exchange Format Framework standards
► Remote Authentication
► Speaker Authentication
► Voice

References

1. Markowitz, J.: The Speaker Verification Application Programmers Interface Standard (SVAPI). In: Harper, D. (ed.) BiometriCon'97: Conference Proceeding. Diane Publishing Company, Darby, PA (1997)
2. Novell Corporation: SRAPI and SVAPI Source Code. (2006). http://developer.novell.com/wiki/index.php/SRAPI_and_SVAPI_Source_Code
3. Colombi, J.: Interface Specification: Human Authentication – Application Program Interface (HA-API) Ver. 2.0. United States Biometrics Consortium, Fort Meade, MD, (1998)
4. Enterprise Computer Technology Forum: S.100 Media Services Volume 6: Media Resources and Services, Revision 2.0 (1998). http://www.comptia.org/sections/ectf/Documents/s100r2v6.pdf
5. Markowitz, J., Rehor, K.: Standards for speaker recognition. In: Proceedings of Biometric Consortium'06, Baltimore, MD (2006)
6. Skerpec, V. (ed.): Speaker Identification and Verification (SIV) Glossary. VoiceXML Forum (2007). http://www.voicexml.org/biometrics/
7. Daboul, C., Eckert, M. (eds.): Speaker Identification and Verification Applications. VoiceXML Forum (2006). http://www.voicexml.org/biometrics/
8. Daboul, C., Shinde, P (eds.): Speaker Identification and Verification (SIV) Requirements for VoiceXML Applications Ver. 2.0. VoiceXML Forum (2007). http://www.voicexml.org/biometrics/
9. INCITS 456, Speaker Recognition Format for Raw Data Interchange (SIVR) (2008). http://www.techstreet.com/incitsgate.tmpl
10. ISO, *ISO 8601 2004(E)* Data Elements and Interchange Formats – Interchange Formats – Representation of Dates and Times. Geneva: International Standards Organization (2004)
11. European Telecommunications Standards Institute: Distributed Speech Recognition; Front-end feature extraction algorithm; Compression algorithms. ETSI document ES 201 108 V1.1.2 2000–04 (2000)

12. Broun, C.C., Campbell, W.M., Pearce, D., Kelleher, H.: Distributed speaker recognition using the ETSI distributed speech recognition standard. In: Proceedings of the International Conference on Artificial Intelligence, pp. 1:244–248 (2001). http://nsodl.org/resource/2200/2006H
13. Oran, D.: Requirements for Distributed Control of Automatic Speech Recognition (ASR), Speaker Identification/Speaker Verification (SI/SV), and Text-to-Speech (TTS) Resources, Internet Informational RFC 4313 (2005). http://www3.tools.ietf.org/html/rfc4313
14. Shanmugham, S., Monaco, P., Eberman, B.: A Media Resource Control Protocol (MRCP) Internet Informational RFC 4463, (2006). http://www.ietf.org/rfc/rfc4463.txt
15. Shanmugham, S., Burnett, D.: Media Resource Control Protocol Version 2 (MRCPv2) (2007) NOTE: This is draft 17. As of December, 2008 it was the current draft. Upon final approval a stable IETF Internet Informational RFC reference number will be assigned. http://tools.ietf.orglid

Speaker Segmentation

Laura Docio-Fernandez, Carmen Garcia-Mateo
Department of Signal Theory and Communications, University of Vigo, Vigo, Spain

Synonyms

Speaker change detection; Speaker clustering; Speaker diarization

Definition

Speaker segmentation is the process of partitioning an input audio stream into acoustically homogeneous segments according to the speaker identity. A typical speaker segmentation system finds potential speaker change points using the audio characteristics.

Introduction

Segmenting an audio–visual stream by its constituent speakers is essential in many application domains. First, for audio–visual documents, speaker changes are often considered natural points around which to structure the document for navigation by listeners (► speaker indexing). In broadcast news, for example, speaker changes typically coincide with story changes or transitions. Audio recordings of meetings, presentations, and panel discussions are also examples where organizing audio segments by speaker identity can provide useful navigational cues to listeners. Furthermore, an accurate speaker segmentation system is also necessary for effective audio content analysis and understanding, audio information retrieval, speaker identification-verification-tracking, and other audio recognition and indexing applications. In fact, speaker segmentation is an important subproblem of the ► speaker diarization task, which is used to answer the question *Who spoke when?*. Speaker segmentation focuses on finding out when a person is speaking and the main goal is to mark where speaker changes occur, i.e., to divide a speech signal into a sequence of speaker-homogeneous regions. Typically, there is no prior knowledge about the speech characteristics of the speakers or the number of different speakers before the process starts, so these have to be derived, in an unsupervised manner, from the same data that are going to be used to find the speaker changing points.

Second, speaker segmentation relates to automatic transcription of speech. In many scenarios, the performance of automatic speech recognition can benefit greatly from speaker adaptation, whether supervised or unsupervised. Speaker segmentation, while not a strict prerequisite for speaker adaptation, is important for performing adaptation on multispeaker data, as it can provide the recognizer with homogeneous speaker data.

Speaker segmentation has sometimes been referred to as speaker change detection and is closely related to acoustic change detection. It has received much attention recently. For a given audio stream, speaker segmentation systems find the times when there is a change of speaker in the audio. On a more general level, acoustic change detection aims at finding the times when there is a change in the acoustics in the recording, which includes speech/nonspeech, music/speech and others. Thus, acoustic change detection can detect boundaries within a speaker turn when the background conditions change.

With the rapid increase in the availability of multimedia data archives, efficient segmentation, indexing and retrieval of audio–visual data is quite an important task in many applications. Automatic metadata extraction from video and audio recordings enables the development of sophisticated multimedia content management applications which can help users manage their

personal recordings. For real world audio–visual data the text can be generated using automatic speech recognition (ASR), the speaker labeled using speaker recognition, and the speaker turns and segments derived can be used for indexing the associated audio and video.

The general unsupervised speaker segmentation problem, in addition to not having models or other information to help segment the speech data by speaker, brings several additional obstacles that complicate the task of separating the segments of one speaker from the segments of another speaker. For example, multispeaker speech data typically includes several short segments. Short segments are difficult to analyze because of the inherent instability of short analysis windows. In addition, more than one speaker may be talking at the same time in multispeaker speech data and the segments may be contaminated with the speech of another speaker. Also, the accuracy of the segmentation process is affected by background noise and/or music. This leads to the need of modeling of these artifacts, which in turn increases system complexity. Other difficulties are related to the dynamic fine-tuning of some parameters that improve the accuracy of the segmentation algorithms. It is also a major concern into optimizing the system performance in terms of access times and signal processing speed. It is highly desirable that these segmentation tasks are accomplished automatically with the least user intervention but additionally these need to be performed fast and accurately.

The task of speaker segmentation can be considered as an evolution of a Voice Activity Detection (VAD), also referred to as Speech Activity Detection (SAD). VAD constitutes a very basic task for most speech-based technologies (Speech Coding, automatic speech recognition (ASR), Speaker Recognition (SR), speaker segmentation, voice recording, noise suppression and others). The classification of an audio recording in speech and nonspeech segments can be utilized to achieve more efficient coding and recognition.

Grouping together segments from the same speaker, i.e., ► speaker clustering, is also a crucial step for segmentation. Speaker segmentation followed by speaker clustering is referred to as *speaker diarization*. Diarization has received much attention recently. It is the process of automatically splitting the audio recording into speaker segments and determining which segments are uttered by the same speaker. In general, diarization can also encompass speaker verification and speaker identification tasks.

Speaker clustering also belongs to the pattern classification family. Clustering data into classes is a well-studied technique for statistical data analysis, with applications in many fields, and, in general, can be defined as unsupervised classification of data, i.e., without any a priori knowledge about the classes or the number of classes. In the speaker diarization task, the clustering process should result, ideally, in a single cluster for every speaker identity. The most common approach is to use a hierarchical agglomerative clustering approach in order to group together segments from the same speaker [1]. Hierarchical agglomerative clustering typically begins with a large number of clusters which are merged pair-wise, until arriving (ideally) at a single cluster per speaker. Since the number of speakers is not known a priori, a threshold on the relative change in cluster distance is used to determine the stopping point (i.e., number of speakers). Determining the number of speakers can be difficult in applications where some speakers speak only during a very short period of time (e.g., in news sound bites or back channels in meetings), since they tend to be clustered in with other speakers. Although there are several parameters to tune in a clustering system, the most crucial is the distance function between clusters, which impacts on the effectiveness of finding small clusters.

Examples of Efforts to Foster Speaker Segmentation Research

The Defense Advanced Research Projects Agency (DARPA) and U.S. National Science Foundation have promoted research in speech technologies for a wide range of tasks from the late 1980s. Additionally, there are significant speech research programs elsewhere in the world, such as European Union funded projects.

The Information Technology Laboratory (ITL) of the National Institute of Standards and Technology (NIST), has the broad mission of supporting U.S. industry, government, and academia by promoting U.S. innovation and industrial competitiveness through advancement of information technology measurement science, standards, and technology in ways that enhance economic security and improve our quality of life.

From 1996 the NIST Speech Group, collaborating with several other Government agencies and research institutions, contributes to the advancement of the state-of-the-art in human language technologies and

related multimodal technologies that employ machine learning approaches by

- Developing measurement methods and algorithms
- Providing annotated corpora for development and evaluation
- Coordinating challenge-task-focused benchmark tests
- Sponsoring evaluation-oriented workshops
- Building test-bed systems

Benchmark tests, implemented within this community since 1987, are used to track the development of several speech technologies. These tests, which provide diagnostic information that helps to identify the strengths and weaknesses of the technology, have facilitated increased accuracy and robustness of the technology over time.

In 1996 NIST also started the 1996 ARPA CSR Hub-4 evaluation (1996–1999). The purpose of this evaluation is to improve the basic performance of speaker-independent unlimited-vocabulary recognition systems using Broadcast News Sources. In this task speaker segmentation enables speaker normalization and adaptation techniques to be used effectively to integrate speech recognition.

In 1997 NIST started the Hub-5E evaluation (1997–2001) that focuses on the task of transcribing conversational speech into text. This task is posed in the context of conversational telephone speech.

Since 1996, NIST has also organized yearly Speaker Recognition (SR) evaluation campaigns, focusing on the automatic ▶ speaker detection and ▶ speaker tracking tasks. In 2000, the NIST SR evaluation introduced the speaker segmentation evaluation as a new task.

With the DARPA EARS (Effective, Affordable, Reusable Speech-to-Text) program (2002–2004) the focus moves on a new task, denoted *rich transcription*, which addresses the need for systems that generate high accuracy, readable transcripts. Here, semantic information is not the only element of interest. Indeed, acoustics-based information (sounds, speech qualities, speaker information, ...), discourse-based information (disfluencies, emotion, ...), as well as linguistic information (topic, named entities, ...) may also be used to enrich the transcription and to help for indexing audio documents. Speaker characteristics are obviously an important information in this context.

The EARS program supports several evaluation tasks that are administrated by the NIST under the Rich Transcription (RT) heading. The specific research tasks are broadly categorized as supporting either Speech-to-Text (STT) or Metadata Extraction (MDE). While STT emphasizes getting the words right, MDE is concerned with structuring STT output to be maximally readable for humans and downstream automatic processes by humans and machines. The Metadata Extraction (MDE) component is designed to enrich the raw word sequence generated by STT systems, by introducing additional information (e.g., who is speaking, how the word stream breaks into sentence units, how to correct the word sequence based on verbal edits) that plays a fundamental role not just in *transcribing* the true speech content but also in facilitating downstream processing by humans and machines.

For this reason, since 2003 the speaker segmentation system evaluation had joined in the Rich Transcription evaluation campaigns and had left the Speaker Recognition evaluation campaign.

The NIST RT metadata MDE task has been including several tracks:

- MDE "Who Spoke When" Speaker Diarization focused on speaker segmentation and clustering.
- MDE "Who Said What" Speaker Diarization.
- MDE Speech Activity Detection.
- MDE Source Localization.
- *Structural* MDE concerned with identifying sentence-like units and detecting disfluencies.

The RT evaluation corpora have included different domains: broadcast news, conversational telephone speech, *conference* room meetings, and *lecture* room meetings. The segmentation challenges are different for the different tasks according to their quality of recordings, number of speakers, the speaking duration of each speaker, and the sequence of speaker changes, etc. But usually high-level speaker segmentation techniques work well over different domains.

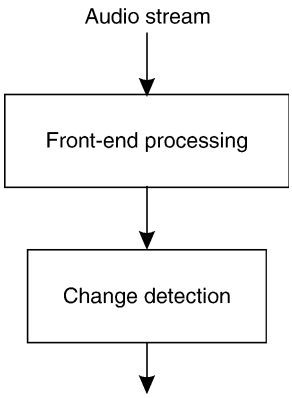
Operation of a Speaker Segmentation System

A basic speaker segmentation system consists of three main steps. First, the input signal is processed to extract a set of acoustic features. Second, a speech/nonspeech detector separates target speech regions from the given audio clip. And lastly, the speaker change detector identifies potential speaker changing

points in each speech region and consequently divides the speech regions into segments containing speech from a single speaker (Fig. 1).

The state-of-art systems for speaker segmentation can be divided into three categories: metric-based, model-based, and hybrid (i.e., combined metric- and model-based) ones. The segmentation process may be carried out by a single pass or by multiple passes through the acoustic data. In the multiple passes case the decision of change-point detection is refined on successive iterations.

Metric-based segmentation is probably the most used approach. It relies on the definition of some metric or distance measure to compare the spectral characteristics on both sides of successive points of the audio signal, and it hypothesizes as speaker change points those boundaries whose distance values exceed a given threshold. The performance of this approach depends highly on the metric and the threshold. Various metrics have been proposed and analyzed in the literature. The most cited are the Bayesian Information Criterion (BIC) which presents the advantages of robustness and threshold independence [2]; the Generalized Likelihood ratio (GLR) and the Kullback–Leibler distance [3]; Divergence Shape Distance [4], etc. The threshold is normally defined empirically given a development set, according to a desired performance. Thus, the threshold will be dependent on the data being processed and needs to be redefined every time data of a different nature need to be processed. This problem has been studied within the speaker identification community in order to classify speakers in an open set speaker identification task [5].

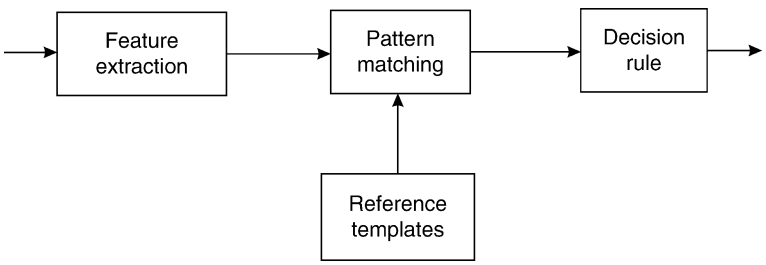


Speaker Segmentation. **Figure 1** A brief flow diagram for a speaker segmentation module.

Model-based techniques are an applied evolution of a common pattern recognition task (Fig. 2). In model-based segmentation, a set of models is estimated for different speaker classes by using training data. Then, the input audio stream is classified, using these models, by finding the most likely sequence of models [6, 7]. The boundaries between models become the segmentation change points. Several models, including Gaussian Mixture Models (GMMs) [8], ▶ Hidden Markov Models (HMMs) [9] and Support Vector Machines (SVMs) [10] have been employed to describe specific speakers.

Hybrid techniques combine metric- and model-based techniques [11]. Usually, metric-based segmentation is used initially to presegment the input audio signal. The obtained segments are used then to create a set of speakers models. Finally, model-based resegmentation gives a refined segmentation.

There are some speaker segmentation techniques proposed in the literature that are not a clear fit to any of the two previous categories. For example, in [12] dynamic programing is proposed to find the speaker change points. In [13] a genetic algorithm is proposed where the number of segments is estimated via the Walsh basis functions and the location of change points



Speaker Segmentation. **Figure 2** Block diagram of a traditional pattern recognition system.

is found using a multipopulation genetic procedure. In [14] segmentation is based on the location estimation of the speakers by using a multiple-microphone setting. The difference between two locations is used as a feature and tracking techniques are employed to estimate the change points of possibly moving speakers.

Assessing Performance

A speaker segmentation system should provide the correct speaker turns and therefore the segments should contain a single speaker. The performance of speaker segmentation can be assessed in terms of the accuracy of speaker turn point detection. In this case, two pairs of figures of merit are commonly used to assess the performance of a speaker segmentation system. On the one hand, one may define two fundamental types of errors, namely false alarm (FA) and missed detection (MD). A FA of turning point detection occurs when a detected turning point is not a true one. A missed MD occurs when a true turning point cannot be detected. Thus, it is possible to use the false alarm rate (FAR) and the miss detection rate (MDR) defined as:

$$\mathrm{FAR} = \frac{N_{\mathrm{FA}}}{N_{\mathrm{FA}} + N_{\mathrm{ref}}},$$

$$\mathrm{MDR} = \frac{N_{\mathrm{MD}}}{N_{\mathrm{ref}}},$$

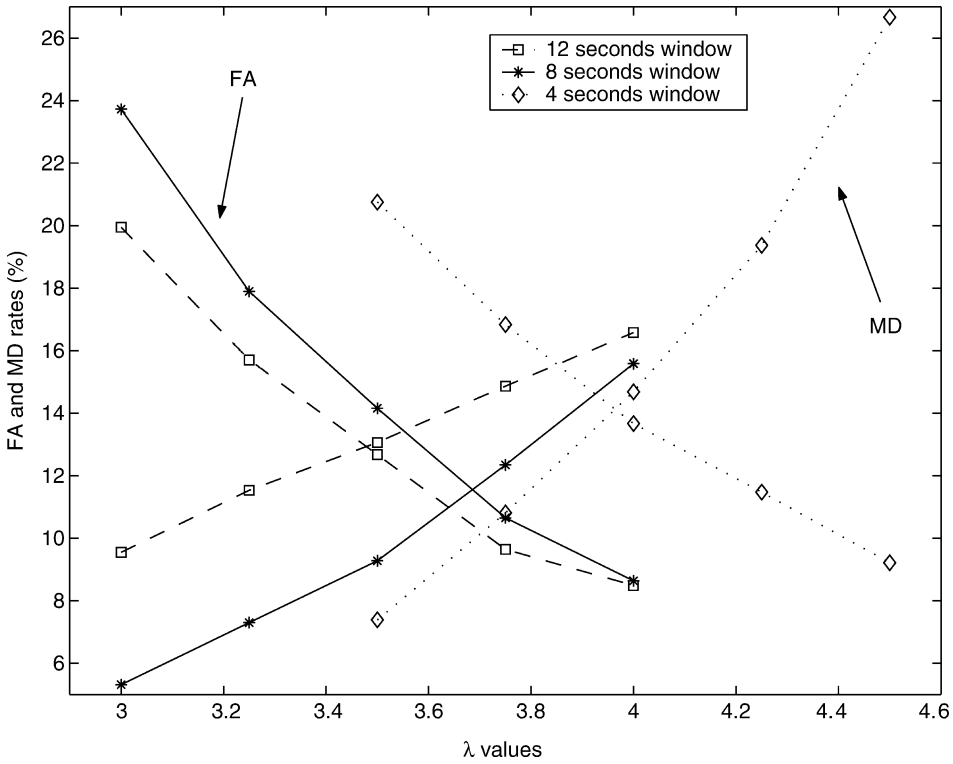
where N_{FA} and N_{MD} are the total number of FA and MD respectively, and N_{ref} is the total number of true turning points given by the reference manual segmentation. A high value of FAR signifies that the speech signal has been oversegmented. A high value of MDR means undersegmentation. Figure 3 shows an example of FAR–MDR curve [15]. This figure shows the tradeoff between missed detection and false alarm, and provides a reference to select different operation points.

On the other hand, one may employ the precision (PRC) and recall (RCL) rates given by

$$\mathrm{PRC} = \frac{\mathrm{CFC}}{\mathrm{DET}},$$

$$\mathrm{RCL} = \frac{\mathrm{CFC}}{\mathrm{GT}},$$

where CFC denotes the number of correctly found changes, DET is the number of the detected speaker changes, and GT stands for the actual number



Speaker Segmentation. Figure 3 The MDR–FAR curve for speaker segmentation.

of speaker turns, i.e., the ground truth. For the latter pair, another objective figure of merit is the F_1 measure

$$F_1 = \frac{2\text{PRCRCL}}{\text{PRC} + \text{RCL}}$$

that admits a value between 0 and 1. The higher its value is, the better performance obtained is. Between the pairs (FAR, MDR) and (PRC, RCL) the following relationships hold:

$$\text{MDR} = 1 - \text{RCL},$$

$$\text{FAR} = \frac{\text{RCLFA}}{\text{DETPRC} + \text{RCLFA}}.$$

The performance of speaker segmentation can also be assessed in terms of the speaker coverage. For the measurement of speaker coverage, the false alarm coverage (FACov) and the missed detection coverage (MDCov) are defined as,

$$\text{MDCov} = \frac{\sum_i \text{duration of missed portion for reference segment i}}{\sum_i \text{duration of reference segment } i},$$

$$\text{FACov} = \frac{\sum_j \text{duration of false portion for detected segment j}}{\sum_i \text{duration of detected segment j}}.$$

In the NIST evaluations the performance of a speaker segmentation system is measured using the segmentation cost function, defined as a weighted sum of decision errors, weighted by error type and integrated over error duration. Thus, five kinds of errors are considered, all as a function of time:

- Missing a segment of speech when speech is present (PMissSeg)
- Falsely declaring a segment of speech when there is no-speech (PFASeg)
- Assigning a false alarm speaker to a segment of speech (PMissSpkr)
- Assigning a speaker to a segment of speech of a missed speaker (PFASpkr)
- Assigning an incorrect speaker to a segment of speech (PErrSpkr)

Therefore, the speaker segmentation cost is defined as:

$$\begin{aligned} C_{\text{Seg}} = {} & (C_{\text{MissSeg}} P_{\text{MissSeg}} + C_{\text{FASeg}} P_{\text{FASeg}}) \\ & + (C_{\text{MissSpkr}} P_{\text{MissSpk}} + C_{\text{FASpk}} P_{\text{FASpk}}) \\ & + C_{\text{ErrSpk}} P_{\text{ErrSpk}}. \end{aligned}$$

Typically, the cost parameters are all set equal to 1.

Applications

First, one of the applications is in multimedia information management (information indexing, information access and content protection) in order to automatically extract meta-data information. Multimedia technologies, which play a crucial role in a wide range of recent application domains, are highly demanded to further facilitate multimedia services and to more efficiently utilize multimedia information generated from diverse domains. A multimedia content based indexing and retrieval system requires analysis of both textual and speaker content. Speaker changes are often considered natural points around which to structure the spoken document for navigation by users. Creating an index into an audio–visual stream, either in real time or in postprocessing, may enable a user to locate particular segments of the audio data. For example, this may enable a user to browse a recording to select audio segments corresponding to a specific speaker, or "fast-forward" through a recording to the next speaker. In addition, knowing the ordering of speakers can also provide content clues about the conversation, or about the context of the conversation. In broadcast news, for example, speaker changes typically coincide with story changes or transitions. Furthermore, audio recordings of meetings, presentations, and panel discussions are also examples where organizing audio segments by speaker identity can provide useful browsing cues to listeners. Also, the audio recording of a meeting or a conversation can be speaker-indexed automatically to facilitate the search and retrieval of the content spoken by a specific person. In this way, meeting information can be obtained conveniently, such as who is saying what and when, remotely through on-line or off-line systems.

Second, biometric applications such as access control. Surveillance is becoming increasingly important for public places. However, most surveillance systems simply store video data, then storing video information selectively through identifying key events or human activities is very important for facilitating access to huge amount of the stored surveillance video archivals with improved browsing and retrieval functionality. Tracking speaker-specific segments in conversations, to aid in surveillance applications, is another place to use a speaker segmentation system.

Third, ASR related applications such as the transcription of conversations. Speaker segmentation relates to automatic labeling and transcription of audio archives

that involve multiple speakers. In this application, the audio signal typically contains speech from different speakers under different acoustic conditions. It is well known that the performance of automatic speech recognition can benefit greatly from speaker adaptation, whether supervised or unsupervised. With the knowledge of "who is speaking," acoustic models for speech recognition can be adapted to better match the environmental conditions and the speakers. Furthermore, in the speech-to-text conversion process, information about speaker turns can also be used to avoid linguistic discontinuity.

Also, capturing the speaker change in a given audio stream could be very useful in military and forensic as well as commercial applications. In forensic applications it is often required to process speech recorded by means of microphones installed in a room where a group of speakers conduct a conversation. Questions such as how many speakers are present, at what time a new person has joined (left) the conversation and others are often asked. It is also often required to determine the true identity of the speakers, or some of them, using available templates of known suspects. For this, one needs to segment the recorded signal into the various speakers and then use conventional speaker identification or verification methods.

Summary

There are a number of relevant applications that may benefit from a speaker segmentation module. Among them, ASR (rich transcription), video tracking, movie analysis, etc. Defining and extracting meaningful characteristics from an audio stream aim at obtaining a more or less structured representation of the audio document, thus facilitating content-based access or search by similarity.

In particular, speaker detection, tracking, clustering as well as *speaker change detection* are key issues in order to provide metadata for multimedia documents and are an essential preprocess stage of multimedia document retrieval. Speaker characteristics, such as the gender, the approximate age, the accent or the identity, are also key indices for the indexing of spoken documents. It is also important information concerning, the presence or not of a given speaker in a document, the speaker changes, the presence of speech from multiple speakers, etc.

Related Entries

- ▶ Gaussian Mixture Models
- ▶ Hidden Markov Models
- ▶ Pattern Recognition
- ▶ Speech Analysis
- ▶ Speaker Features
- ▶ Session Effects on Speaker Modeling
- ▶ Speaker Recognition, Overview

References

1. Docio-Fernandez, L., Garcia-Mateo, C.: Speaker segmentation, detection and tracking in multi-speaker long audio recordings. In: Third COST 275 Workshop: Biometrics on the Internet, pp. 97–100. Hatfield, UK (2005)
2. Chen, S.S., Gopalakrishnan, P.: Clustering via the bayesian information criterion with applications in speech recognition. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, pp. 645–648. Seattle, WA (1998)
3. Delacourt, P., Wellekens, C.J.: DISTBIC: a speaker-based segmentation for audio data indexing. Speech Commun. **32**(1–2), 111–126 (2000)
4. Lu, L., Zhang, H.J.: Speaker change detection and tracking in real-time news broadcasting analysis. In: ACM International Conference on Multimedia, pp. 602–610. Quebec, QC, Canada (2002)
5. Campbell, J.P.: Speaker recognition: a tutorial. Proc. IEEE **85**(9), 1437–1462 (1997)
6. Gauvain, J.L., Lamel, L., Adda, G.: Partitioning and transcription of broadcast news data. In: Proceedings of International Conference on Speech and Language Processing, vol. 4, pp. 1335–1338. Sidney, Australia (1998)
7. Kemp, T., Schmidt, M., Westphal, M., Waibel, A.: Strategies for automatic segmentation of audio data. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1423–1426. Istanbul, Turkey (2000)
8. Gauvain, J.L., Lamel, L., Adda, G.: The LIMSI broadcast news transcription system. Speech Commun. **37**(1–2), 89–108 (2002)
9. Moraru, D., Meignier, S., Fredouille, C., Besacier, L., Bonastre, J.F.: The ELISA consortium approaches in broadcast news speaker segmentation during the NIST 2003 rich transcription evaluation. In: Proceedings of IEEE ICASSP'04, pp. 223–228. Montreal, Canada (2004)
10. Lu, L., Li, S.Z., Zhang, H.J.: Content-based audio segmentation using support vector machines. ACM Multimedia Syst. J. **8**(6), 482–492 (2001)
11. Kim, H.G., Ertelt, D., Sikora, T.: Hybrid speaker-based segmentation system using model-level clustering. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 1, pp. 745–748. Philadelphia, PA (2005)
12. Vescovi, M., Cettolo, M., Rizzi, R.: A DP algoritm for speaker change detection. In: Proceedings of Eurospeech03. (2003)

13. Pwint, M., Sattar, F.: A segmentation method for noisy speech using genetic algorithm. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Philadelphia, PA (2005)
14. Lathoud, G., McCowan, I., Odobez, J.: Unsupervised location-based segmentation of multi-party speech. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing: NIST Meeting Recognition Workshop. Montreal, Canada (2004)
15. Perez-Freire, L., Garcia-Mateo, C.: A multimedia approach for audio segmentation in TV broadcast news. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 369–372. Montreal, QC, Canada (2004)

Speaker Separation

Speaker separation is a technology used in multi-speaker environments to separate the vocal features of each speaker from those of the other speakers, even when speakers interrupt and talk over each other.

► Speaker Recognition, Standardization

Speaker Tracking

Speaker tracking consists of determining not only whether a particular speaker appears in a multispeaker audio stream, but identifying the specific intervals within the audio stream corresponding to the speaker. It requires that this speaker is known a priori by the system. In that sense, speaker tracking can be seen as a speaker verification task applied locally along a document containing multiple interventions of various speakers. The objective of this task is to cluster the speech by speaker.

► Speaker Segmentation

Speaker Verification

► Liveness Assurance in Voice Authentication

Spectral Analysis of Skin

► Skin Spectroscopy

Specular Reflection

Specular reflection is the mirror-like reflection of light or waves on a surface. The incoming light is reflected at the same angle as it hits on the surface.

► Iris Standards Progression
► Skin Spectroscopy

Specularity

► Specular Reflection

Speech Analysis

Doroteo T. Toledano, Daniel Ramos, Javier Gonzalez-Dominguez, Joaquín González-Rodríguez
ATVS – Biometric Recognition Group. Escuela Politecnica Superior, Universidad Autonoma de Madrid, Spain

Synonyms

Speech parametrization

Definition

The analysis of speech signals can be defined as the process of extracting relevant information from the speech signal (i.e., from a recording). This process is mainly based on the speech production mechanism, whose study involves multiple disciplines from linguistics and articulatory phonetics to signal processing and

source coding. In this article, a short overview is given about how the speech signal is produced and typical models of the speech production system, focusing on the different sources of individuality that will be present in the final uttered speech. In this way, the speaker who produced the speech with those individual features is then recognizable both for humans and for machines.

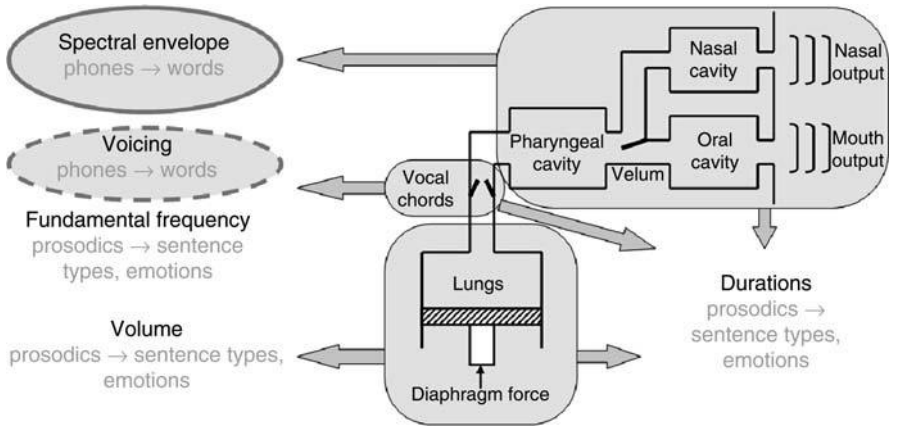
Although speech production is felt by humans as a very natural and simple mechanism, it is a very complex process that involves the coordinated participation of several physiological structures that evolution has developed over the years. For a deeper description of this process the interested reader may consult some of these excellent books [1–3]. Here the human speech production mechanism is described very briefly as the basis for the automated speech analysis systems. Once these mechanisms have been understood, the most common methods to analyze speech are addressed. These methods are based on the speech production mechanisms to some extent. The last part of this article analyzes how the relevant information in this context (the speaker individualization information) is encoded into the speech signal.

Speech Production and Its Relation to Speech Analysis

The process of speech production is described in many books [1–3]. Here the main conclusions about how the speech production system relates to the main parameters estimated in speech analysis are briefly reviewed. Next section addresses the problem of estimating these parameters.

Figure 1 summarizes the different parts of the human speech productionsystem (represented very schematically), and how they are related to the main parameters that describe the characteristics of the speech signal:

- *Volume or intensity of the sound.* The volume or intensity of a speech sound depends mainly on the amount of air exhaled by the lungs and the muscular tension on the articulators producing the sound. The volume or intensity is a prosodic parameter that is related to emotions (i.e., speech in anger has usually more volume than normal or relaxed speech) and sentence type (for instance interrogative sentences tend to end with a higher intensity).
- *Voicing* and, in case of a voiced sound, ▶ fundamental frequency. Human sounds can be voiced or unvoiced depending on whether the vocal chords vibrate or not when producing the sound. Voicing is a binary feature that is essential in discriminating different phonemes. In voiced sounds the vocal chords vibrate at a frequency that is called fundamental frequency (also called F0, tone, or pitch). The fundamental frequency depends on the tension applied to the vocal chords and on the air flow produced by the lungs, and can be modulated to provide the sentence with a certain intonation, constituting one of the most important prosodic parameters. The fundamental frequency plays an important role in determining emotions and sentence types.



Speech Analysis. Figure 1 Simplified functional scheme of the human speech production system with indication of the speech parameters affected by each organ.

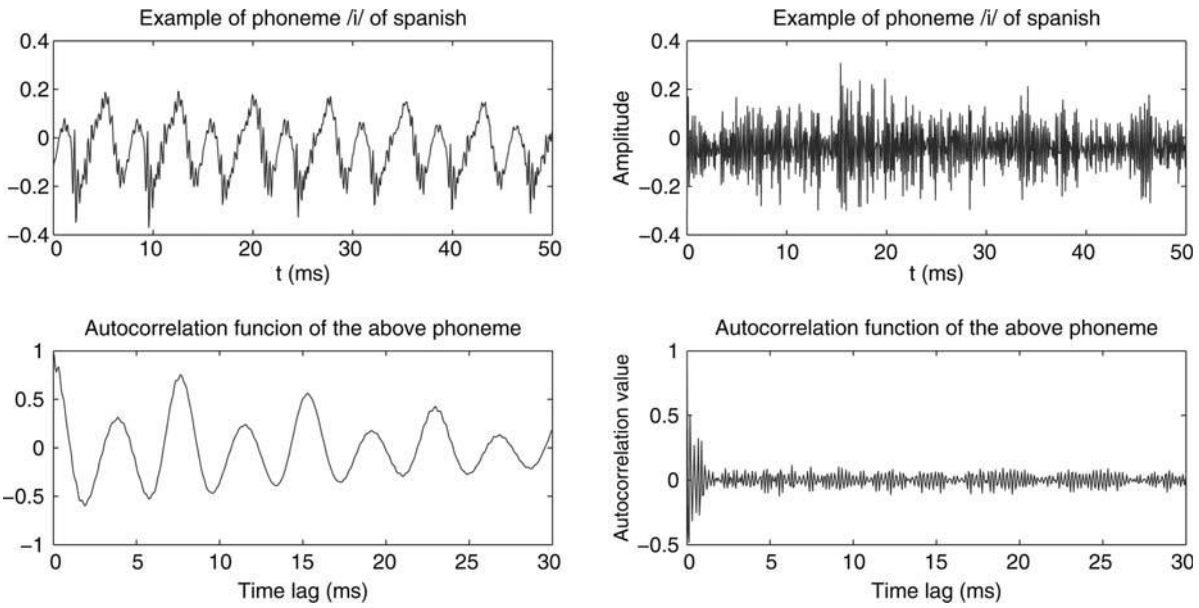
- *Spectral envelope.* The rest of the speech production system from the vocal chords to the lips and nostrils is called the *vocal tract*. The effect of the vocal tract is to modulate the sound produced to obtain the different phonemes. This is accomplished with the help of several mobile parts called *articulators* such as the tongue, the lips, and the teeth, which can substantially modify the shape of the vocal tract and therefore the modulation produced. It can be seen that the modulation produced by the vocal tract affects mainly the spectral envelope of the signal produced. This spectral envelope typically presents a few maxima at the frequencies of resonance of the vocal tract, called *formants*, which are characteristics of the different phonemes. In fact, it is possible to distinguish the different vowels based on their formants. The spectral envelope alone is capable of, with the help of voicing (and fundamental frequency for tonal languages such as Chinese), discriminating among the different phonemes of a language and also among different speakers.
- *Duration of the phonemes.* The speech production system moves over time in a coordinated way and this movement defines the durations of the phonemes. This is considered a prosodic feature that contains valuable information for recognizing phonemes and speakers.

This complex system is coordinated and directed by the brain, which in a much more complex and largely unexplained process is capable of generating the adequate sequence of words to utter at a precise instant in a dialog, transforming these sentences into a sequence of phonemes, sending the necessary orders to the muscles to coordinately produce the speech and even superimposing other information such as emotions. This process of language generation is mainly learned, and different individuals learn to generate language and coordinate the articulatory organs in different ways, thus constituting another source of speaker discriminating information.

Speech Analysis

Speech analysis is the process of analyzing the speech signal to obtain relevant information of the signal in a more compact form than the speech signal itself. Given the previous review of the speech production mechanism and its relation to the most important characteristics of speech, the goal of speech analysis is to obtain some or all of these parameters (and possibly more) from a speech recording. This section presents a review of how these parameters are estimated from a speech recording and how important they are for voice biometrics.

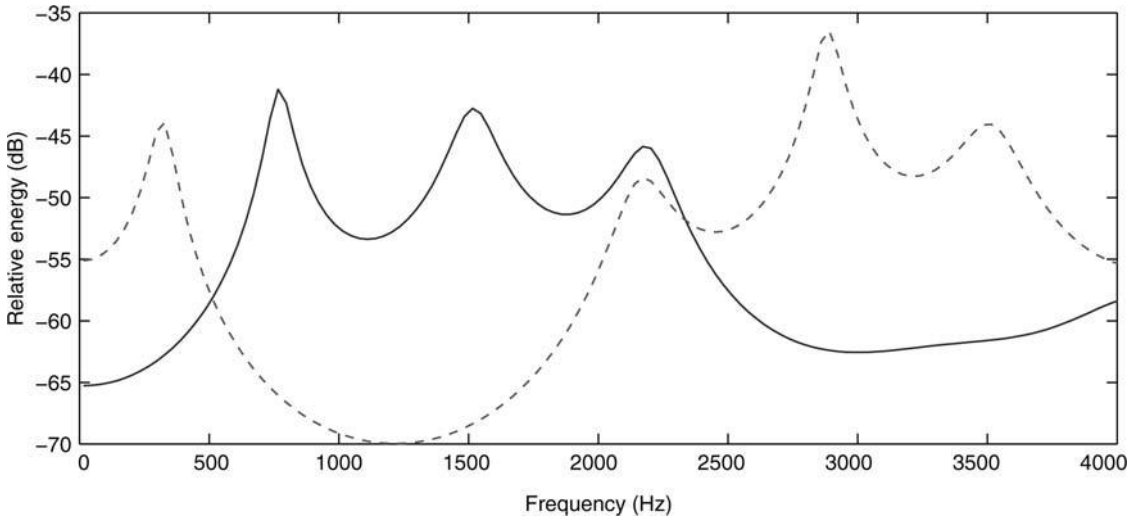
- *Volume or intensity of the sound.* This parameter is typically measured as the logarithm of the short-term average energy of the signal (i.e., the average of the energy of the signal over a few milliseconds). Intensity can be a clue to identify a speaker and to discriminate between sounds, but this feature is affected very much by external parameters such as the gain of the recording equipment and microphone and even the distance and position between the mouth and the microphone. For this reason absolute intensity is rarely used in speech analysis and only relative intensity variations are used.
- *Voicing* and, in case of a voiced sound, fundamental frequency. Voicing and the fundamental frequency can be estimated from the autocorrelation function of the speech signal. Figure 2 shows a voiced and and unvoiced phoneme and their autocorrelation functions. The quasi-periodicity of the voiced signals becomes apparent in the autocorrelation function as a local maximum at a lag corresponding to the pitch period (the inverse of the fundamental frequency). In Fig. 2 this maximum is placed at a lag of 7.5 ms, corresponding to a fundamental frequency of 133 Hz. For unvoiced phonemes this maximum does not appear. To estimate the fundamental frequency it is necessary to locate the *correct* local maximum in the autocorrelation function, which is sometimes (as in the example shown here) difficult due to the presence of local maxima at rational multiples of the fundamental frequency. Besides the autocorrelation method there are other methods to estimate the fundamental frequency, either with lower computational cost (such as using the *Magnitude Difference Function* instead of the autocorrelation function) or with more precision [4]. The fundamental frequency is very characteristic of the speaker and is very different for male and female speakers. The evolution of the fundamental frequency over time determines the



Speech Analysis. **Figure 2** Example of a voiced sound /a/ and an unvoiced sound /s/ of Spanish and their corresponding autocorrelation functions showing a possible way of determining voicing and estimating the pitch period.

intonation of the utterance and intonation is also very characteristic of the speaker.

- *Spectral envelope*. The spectral envelope of the speech signal contains the richest information about the speech sounds and also about the speaker. Not surprisingly, speech and speaker recognition systems typically focus primarily (many times exclusively) on extracting and processing this dynamically changing information from the speech signal. For this reason many times the speech analysis phase is reduced to the spectral envelope estimation. Several modeling strategies have been proposed in the literature, but with no doubt the most successful one in terms of number of applications based on one kind of modeling is Linear Predictive Coding (LPC) of speech. In this approach, the vocal tract is modeled as an all-pole (or autoregressive, AR) model [5] representing the vocal tract resonances with a digital filter completely determined by the poles positions. In this way, with a very small number of LPC coefficients (typically between 10 and 20), the spectral envelope is fully determined for every analysis frequency. An example of this analysis is shown in Fig. 3 for two different vowels in Spanish. Here the spectral envelope is represented with 17 LPC coefficients and shows very clearly the different formants of the two vowels. Theoretically, only nasal sounds are not properly modeled with an AR system, as the nasal cavity in parallel, as shown in the acoustic theory of tubes, introduces zeros (minima, as opposed to the poles or resonances) in the overall vocal tract response. Then those nasal sounds could be better modeled by ARMA (AR and MA, moving average) – or pole-zero models – doubling the number of coefficients. However, good enough approximations of nasal spectral envelopes can be obtained with the typical number of poles (now in positions not so well correlated with physical configuration of the acoustic vocal tract), simplifying the model with a single all-pole model for all kinds of sounds. Each possible sound is then modeled as a set of LPC coefficients (see "Speaker Features" entry for details on LPC coefficients computation). In this LPC context, the speech production system, as a generator of a continuum of different sounds which constitute syllables, words, and phrases, is modeled as a discrete sequence of different configurations of the LPC model, switching every new analysis frame (typically, each 5–25 ms) to a new vector of parameters defining the model characteristics. This kind of modeling has been successful in



Speech Analysis. **Figure 3** Spectral envelope of /a/ (*solid blue line*) and /i/ (*dashed red line*) Spanish vowels estimated with LPC analysis of order 17 on 8 kHz bandwidth speech (only 0–4 kHz range is shown). The spectral envelope shows clearly the different position of the formants in both vowels.

many applications, such as coding or recognition of speech signals. In speech coding, basic vocoders were based mainly in the model description mentioned earlier, focused on efficient extraction from real speech of the best set of model parameters (also including voicing, fundamental frequency, and intensity) that better fit the actual speech in each analysis frame. Newest codecs have based their improvements in better modeling of the excitation signal, as having catalogs (VQ, Vector Quantized codebooks) of possible excitation patterns, but the underlying model is basically the same as mentioned earlier. In speech recognition, the objective is to properly estimate the phone which better corresponds to the observed spectral features at the input at every time frame. In order to have an efficient (both in accuracy and complexity) pattern recognizer, the coefficients in the feature vector to be modeled should not be correlated, which eases the obtention of pseudo-diagonal covariance matrices modeling the underlying data classes. This is the main reason why cepstral derived features (see "Speaker Features" entry for details) are preferred from highly correlated LPC coefficients, but the model is still valid as the objective is to better decode the phoneme at the origin of the observed (LPC or cepstral) feature vector. Finally, in voice biometrics even this simple LPC model can provide speaker specific information as frame based spectral, but also phonotactic data can be derived from the basic previous features. Details on state-of-the-art features and models for voice biometrics are detailed in corresponding entries "Speaker Features" and "Speaker Modeling."

- *Duration of the phonemes.* Estimating the durations of the phonemes requires recognizing the phonemes and determining the boundaries between them. This process is usually made within the context of phonetic recognition and is generally considered too complex and not enough reliable to be used in the context of speaker recognition, even though durations contain important information for speaker individualization purposes.

Speaker Information in the Speech Signal

Speech production is an extremely complex process that encodes multiple types of information into a speech signal. This section describes the information about the speaker that is encoded in the speech signal. This information is what it is necessary to extract from the speech signal for performing speaker recognition. There is no single way of looking for speaker information in a speech signal. Rather, there are multiple ways

of extracting valuable speaker information from different levels of the speech signal. Recently these levels have been named high-level and low-level speaker features, however, there is more of a continuous rather than a hard division. Some of the levels from which it is possible to extract information about the speaker from a speech signal are the following:

- *Idiolectal characteristics* of a speaker's speech are on the highest levels to take into account, and describe how a speaker uses a specific linguistic system. This "use" of the language is basically learned and is determined by how the speaker learned to generate the adequate words for each speaking act. It can be seen that there are individualities in this use that can be exploited for voice biometrics.
- *Phonotactics* describes the use by the speaker of the phone sequences, highly influenced by the language being spoken but including highly individualized features. A bit lower than the idiolectal characteristics, the phonotactics is also learned by the speaker and determine the phones produced for a sequence of words. As with idiolectal characteristics, it has also been shown that this information has important individualization power.
- *Prosody* is the combination of instantaneous energy, fundamental frequency, and phoneme durations to provide speech with naturalness and full sense. Prosody helps clarifying the message, the type of sentence, and even the state of mind of the speaker. Some prosodic features are learned by the speaker (such as the different prosodic structures for the different messages and possibly even state of mind), but some other prosodic features have a physiological basis (such as the average fundamental frequency). In both cases the prosodic features provide useful speaker information for voice biometrics.
- *Short-term spectral characteristics* are the lowest level features containing speaker individualization information. These are directly related to the articulatory actions related to each phone being produced. Spectral information intends to extract the peculiarities of speaker's vocal tracts and their respective articulation dynamics. Again these features are a mixture of learnt uses (such as dynamics) and physiological features (such as the length of the vocal tract, that have a strong impact on the characteristics of the produced speech).

Summary

An overview of the speech production system has been given, centered on the basic mechanisms involved in speech production and the origin of sounds or phonemes individuality, which makes them recognizable. But in this homogenizing environment (the use of a common linguistic system, usually a language, intended for communication based in common elements), speakers introduce individual characteristics making each speaker's speech to sound according to his individual physical, emotional, and idiolectal characteristics. Simple analysis models as Linear Predictive Coding of speech allow us to easily understand the potential of digital signal processing and pattern recognition techniques, which will lately allow us to build efficient speech codecs or recognizers and even finally good detectors of individual speaker's voice.

Related Entries

- ▶ Session Effects on Speaker Modeling
- ▶ Speaker Features
- ▶ Speech Production
- ▶ Voice, Forensic Evidence of
- ▶ Voice Device

References

1. Huang, X., Acero, A., Hon, H.W.: Spoken Language Processing. Prentice Hall PTR, Upper Saddle River, NJ (2001)
2. Rabiner, L., Schafer, R.: Digital Processing of Speech Signals. Prentice Hall, Upper Saddle River, NJ (1978)
3. Deller, J., Hansen, J., Proakis, J.: Discrete-Time Processing of Speech Signals, 2nd edn. Wiley, New York (1999)
4. Chu, W.C.: Speech Coding Algorithms. Foundation and Evolution of Standardized Coders. Wiley, New York (2003)
5. Oppenheim, A., Schafer, R., Buck, J.: Discrete-Time Signal Processing. 2nd ed. Prentice Hall, Upper Saddle River, NJ (1999)

Speech Input Device

▶ Voice Device

Speech Parametrization

► Speech Analysis

Speech Processing

Speech processing is a technology that operates on the stream of speech.

► Speaker Recognition, Standardization

Speech Production

Laura Docio-Fernandez, Carmen Garcia-Mateo
University of Vigo, Vigo, Spain

Synonyms

Speech system; Sound generation

Definition

Speech production is the process of uttering articulated sounds or words, i.e., how humans generate meaningful speech. It is a complex feedback process in which also hearing, perception, and information processing in the nervous system and the brain is involved.

Speaking is in essence the by-product of a necessary bodily process, the expulsion from the lungs of air charged with carbon dioxide after it has fulfilled its function in respiration. Most of the time one breathes out silently; but it is possible, by contracting and relaxing the vocal tract to change the characteristics of the air expelled from the lungs.

Introduction

Speech is one of the most *natural* forms of communication for human beings. Researchers in speech technology are working on developing systems with the ability to understand speech and speak with a human being.

Human–computer interaction is a discipline concerned with the design, evaluation, and implementation of the most natural interactive computing systems for human use [1]. Computers with this kind of ability are gradually becoming a reality today, through the success of speech synthesis, speech recognition, and other related speech technologies. However, in order to give them functions that are much closer to those of human beings, one must learn more about the mechanisms by which speech is produced and perceived, and develop speech information processing technologies that make use of these functions.

However, progress in advanced computer speech interfaces is limited in part due to incomplete knowledge of the physics of speech production. For computer generated speech output, this means limitations in the naturalness and intelligibility of synthetic speech.

The generation of human speech involves a remarkably complex process. In modeling the process of human speech production one may recognize two principal stages:

1. Formation in the mind of thoughts to be expressed as well as the choice of words to be used. The message is organized on the linguistic level and structured grammatically and phonologically.
2. The string of phonemes is converted into a set of continuous signals controlling the musculature of the various articulators. This results in a highly complex integrated movement sequence in which generally participate all the articulators, the lips, the tongue, the mandible, etc. Finally, the physical interaction of the vibrating vocal cords and the moving articulatory structure produces a continuous acoustic signal perceived as speech.

Speech production is an activity embodied in a complex physical system. It is produced by a cooperation of lungs, glottis (with vocal cords), and articulation tract (mouth and nose cavity). The speaker produces a *speech signal* in the form of pressure waves that travel from the speaker's head to the listener's ears. This signal consists of variations in pressure as a function of time and is usually measured directly in front of the mouth, the primary sound source. The amplitude variations correspond to deviations from atmospheric pressure caused by traveling waves.

An audible speech signal is produced by moving the vocal articulators to modify and "sculpt" the source of sound energy in the vocal tract (e.g., air turbulence or vocal fold vibration). The signal is *non-stationary* changing characteristics as the muscles of the vocal tract contract and relax. Since speaker wishes to produce a ► sound sequence corresponding to the message to be conveyed, most major vocal tract movements have a voluntary basis. For each sound, there is a positioning for each of the vocal tract articulators: vocal cords, tongue, lips, teeth, velum, and jaw. Sounds are typically divided into two broad classes: *vowels*, which allow unrestricted airflow in the vocal tract; and *consonants*, which restrict airflow at some point and have a weaker intensity than vowels.

The most common sound generation sources are quasi-periodic vibration of the vocal cords and turbulent noise generated by the passage of air through a narrow constriction, usually in the oral cavity. More rarely, sounds are generated by plosive release of air (following the buildup of pressure behind an obstruction in the vocal tract), implosion (following the creation of a vacuum behind an obstruction in the vocal tract), and clicks created by, the action of the tongue pulling away from the roof of the mouth.

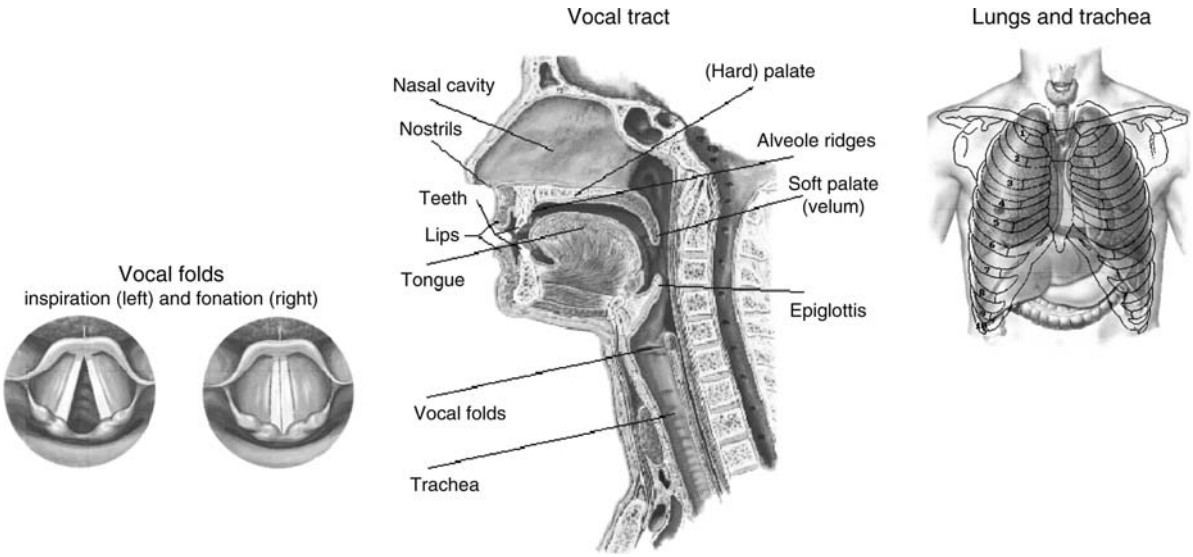
Depending on the type of excitation, i.e., sound generation sources, two types of sounds are produced: voiced and unvoiced sounds. ► Voiced sounds are produced by forcing air through the glottis or an opening between the vocal folds. Then the vocal folds vibrate, they interrupt the air stream and produce a quasi-periodic pressure wave that excite the vocal tract. An example of voiced sound is the vowel "a" in cut, or "ee" in "beet." ► Unvoiced sounds are generated by forming a constriction at some point along the vocal tract and forcing air through the constriction to produce turbulence. Vocal folds do not vibrate in this case. An example of unvoiced sound is "s" as in "six."

The frequency at which vocal folds open and close is called the *fundamental frequency.*

A major focus of speech production research is in modeling articulatory–acoustic relationships of speech sounds. Physically and physiologically based models for speech acoustics are particularly important for developing high-quality speech synthesis and low bit rate (articulatory) coding. Significant progress has been recently made towards developing improved articulatory-acoustic models.

The Human Speech Production Mechanism

Figure 1 shows an illustration of the human speech production system. The gross anatomical components of the systems are the lungs, trachea, larynx (organ of speech production), pharyngeal cavity (throat), buccal cavity (mouth), and nasal cavity (nose).



Speech Production. Figure 1 Schematic view of human speech production mechanism.

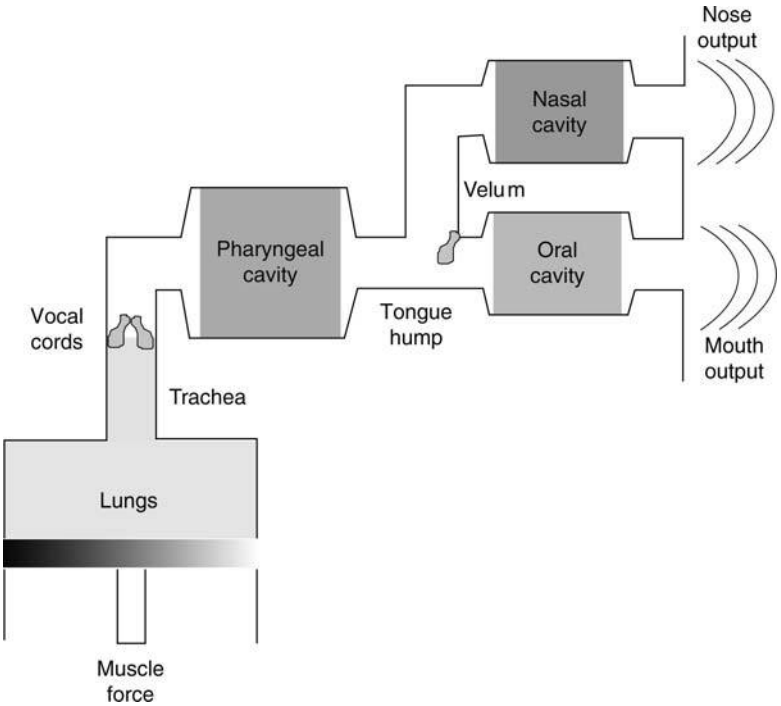
The pharyngeal and buccal cavities are usually grouped and referred to as the *vocal tract*, and the nasal cavity is often called *nasal tract*. Accordingly, the vocal tract begins at the output of the larynx, and terminates at the input of the lips. The nasal tract begins at the velum and ends at the nostrils of the nose. There are many other anatomical components that contribute to the production of speech, such as the vocal folds (or cords), tongue, lips, teeth, and jaw. These are referred to as *articulators* and move to different positions in order to produce various speech sounds. Due to the physical constraints of the vocal tract, the positions of the articulators can only change slowly with time and individual realizations of a phone are strongly influenced by previous and future phones in an utterance. This phenomenon is known as *coarticulation* and is important for both accurate speech analysis and natural speech production.

The speaker-specific characteristics of speech are due to differences in physiological and behavioral aspects of the speech production system in humans. The main physiological aspect of the human speech production system is the vocal tract shape.

The process of human speech production can be summarized as follows. While speaking, the air pushed out from the lungs (the main energy source) travels into the trachea, then up into the glottis, where it is periodically interrupted by the movement of the vocal cords. The tension of the vocal cords is adjusted by the larynx so that the chords vibrate in an oscillatory fashion, resulting in the production of voiced speech. During unvoiced speech, constrictions within the vocal tract (oral cavities – mouth, throat, etc.) force air through the constriction to produce turbulence.

Speech production can be viewed as a filtering operation (source-filter model of speech production [2]) in which the three main cavities of the speech production system (vocal and nasal tracts) comprise the main acoustic filter. The filter is excited by the organs below it (▶ glottal excitation), and is loaded at its main output by a radiation impedance due to the lips (▶ lip-radiation effect). The articulators, most of which are associated with the filter itself, are used to change the properties of the system, its form of excitation, and its output loading over time. Figure 2 shows this model. The source which excites the filter may be either periodic, resulting in voiced speech, or noisy and aperiodic, causing unvoicing speech.

The basic assumption of the model is that the source signal produced at the glottal level is linearly filtered



Speech Production. **Figure 2** Block diagram of human speech production system.

through the vocal tract. The resulting sound is emitted to the surrounding air through radiation loading (lips). The model assumes that source and filter are independent of each other. Although recent findings show some interaction between the vocal tract and a glottal source [3], Fant's theory of speech production is still used as a framework for the description of the human voice.

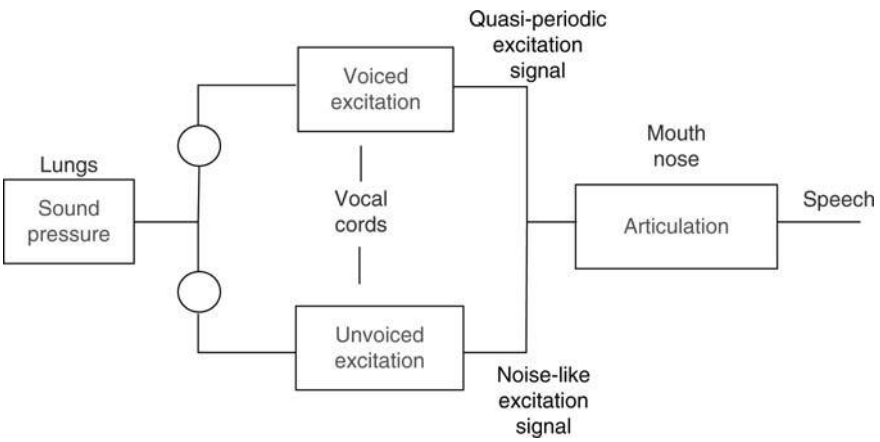
An articulatory representation of the speech production system has certain attractive properties which might be exploited to help in modeling speech. Articulation is the term used for all actions of the organs of the vocal tract that effect modifications of the signal generated by the voice source. This modification results in speech events which can be identified as vowels, consonants or other phonological units of a language. Speech articulators move relatively slowly and smoothly, and their movements are continuous. The mouth cannot jump instantaneously from one configuration to a completely different one. Using speech production knowledge could help to improve speech processing methods by providing useful constraints. Suggested applications include, for example, automatic speech recognition, low bit-rate speech coding, speech analysis and synthesis, and animated talking heads.

Modeling the Human Speech Production System

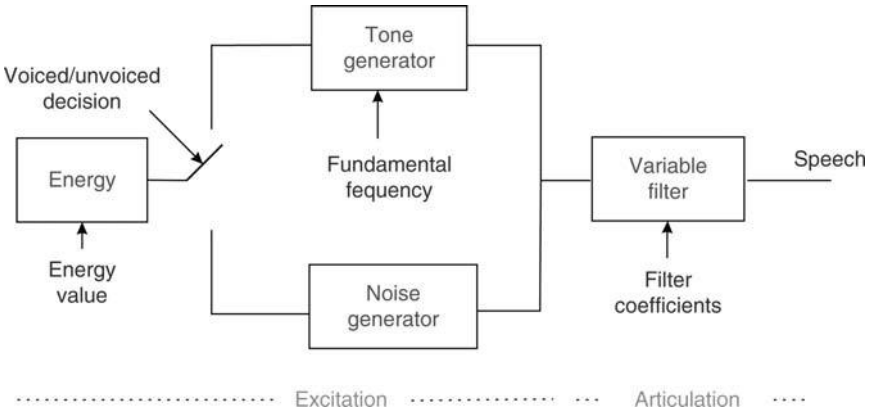
Speech is transmitted between humans in the acoustic domain, and fortunately, it can be easily measured and recorded as acoustic representation. Underlyingly, however, the acoustic speech signal is the product of events in a speaker's articulatory system, and there has long been interest in ways to exploit the underlying articulatory information for speech technology. Increasingly detailed and sophisticated models about how speech is generated in human speech production system has been developed in the past 30 years [4].

The most relevant models of the speech production mechanism belong to acoustic modeling, that is based on the acoustic theory of speech production [5]. According to this theory, speech waveform is considered to be the output of a resonant network (namely the vocal tract filter) that is excited by sound sources placed at the glottis. The main sections of the speech production mechanism, namely the voice source, vocal tract, and radiation effects, are likely to be linearly modeled in a noncoupled manner following a source-filter arrangement. The assumption that the source and the filter can be separately modeled probably holds for most of the cases. However, this assumption is questionable for low frequencies, because the nonlinear coupling may produce damping of the first formant. It is also disputable for unvoiced speech (excitation is due to turbulence originating at constrictions on the vocal tract itself).

As described in the previous section, the human speech production can be illustrated by a simple source-filter model (Fig. 3). Here the lungs are replaced by a DC source, the vocal cords by an impulse generator, and the articulation tract by a linear filter system. A noise generator produces the unvoiced excitation. In practice, all sounds have a mixed excitation, which



Speech Production. Figure 3 A general model for speech production.



Speech Production. **Figure 4** A simplified discrete time model for speech production.

means that the excitation consists of voiced and unvoiced portions. Of course, the relation between these portions varies strongly with the sound being generated.

In general, the source-filter model is related to linear prediction. Based on this model, a further simplification can be made (Fig. 4). A "hard" switch which only selects between voiced and unvoiced excitation is used. The filter, representing the articulation tract, is a simple recursive digital filter; its resonance behavior (frequency response) is defined by a set of filter coefficients. Since the computation of the coefficients is based on the mathematical optimization procedure of Linear Prediction Coding, they are called Linear Prediction Coding Coefficients or LPC coefficients and the complete model is the so-called LPC Vocoder (Vocoder is a concatenation of the terms 'voice' and 'coding') [6].

Speech can be modeled as the response of linear time varying system with appropriate excitation. For voiced speech, the excitation can be approximated by a pulse train in which the pulses appear according to the instantaneous pitch rate. If a single pitch period is analyzed at a time, an analysis known as "pitch synchronous analysis," only one pulse occurs somewhere in the period.

This model has a great advantage. Since the main parameters of the speech production, namely the pitch and the articulation characteristics, expressed by the LPC coefficients, are directly accessible, the audible voice characteristics can be widely influenced. Also the number of filter coefficients can be varied to influence the sound characteristics, above all, the formant characteristics.

Articulatory models of speech production mechanisms aim at modeling the physical, anatomical, and physiological functioning of the organs involved in human voice production. In this approach, the system is modeled instead of the signal or its acoustic characteristics. Modeling the process at the articulatory level can be expected to be simpler because the articulators respond to muscular forces with predictable changes in their position and rates of movement.

It is worth noting that, globally, the speech signal is a nonstationary signal. Then, all the systems in such speech production models should be time-varying, with their parameters changing in accordance to the sound to be produced. The classic source-filter linear model is satisfactory only as a first approximation of the overall nonlinear process of speech production, and only for short time frames, on which the signal is quasi-stationary.

The variability of the speech signal originates in the specific dynamics of the articulatory apparatus. It is well-known that the phonatory system is a time-varying system, and consequently speech signal is nonstationary. A large class of nonstationary and nonlinear processes is involved in speech production.

Applications

To provide a compact computational model for speech production that can be beneficial to a wide range of areas in speech signal processing.

The speech production models which have been derived can be applied in almost all fields of speech processing like speech synthesis, speech analysis, speech and speaker recognition, and also in speech coding.

Summary

Decades of research have gone into the technical and quantitative understanding of human speech production mechanisms.

This communication focuses on the human speech production mechanism. Speech is produced through the careful movement and positioning of the vocal-tract articulators in response to an excitation signal that may be periodic at the glottis, or noiselike due to a major constriction along the vocal tract, or a combination.

Related Entries

- ▶ Biometrics, Overview
- ▶ Speaker Features
- ▶ Speech Analysis

References

1. Hewett, T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., Verplank, W.: Chapter 2: Human–computer interaction. In: B. Hefley (ed.) ACM SIGCHI Curricula for Human–Computer Interaction. ACM, (2007)
2. Fant, G.: Acoustic Theory of Speech Production, 1st edn. Mouton, The Hague (1960)
3. Fant, G.: Glottal flow: models and interaction. J. Phon. **14**, 393–399 (1986)
4. Kent, R.D., Adams, S.G., Turner, G.S.: Models of speech production. In: Principles of Experimental Phonetics, pp. 2–45. N.J. Lass, Mosby (1996)
5. Burrows, T.L.: Speech Processing with Linear and Neural Network Models (1996)
6. Deller, J.R., Proakis, j.G., Hansen, J.H.L.: Discrete-Time Processing of Speech Signals, 1st edn. Macmillan, New York (1993)

Speech Recognition

Speech recognition is a non-biometric technology that enables a machine to recognize the words a speaker speaks.

- ▶ Remote Authentication
- ▶ Speaker Recognition, Standardization

Speech Spectral Envelope

The speech spectral envelope is referred as an approximation to the frequency response of the vocal tract obtained from the speech signal. The speech spectral envelope contains the most discriminative information available in speech to distinguish phonemes and speakers.

- ▶ Speech Analysis

Speech Synthesis

- ▶ Voice Sample Synthesis

Speech System

- ▶ Speech Production

Speed

- ▶ Signature Recognition

SPME

SPME is a simple, solvent-free headspace extraction technique in which volatile and semi-volatile organic compounds adhere to the fiber either by adsorption or absorption depending on the fiber type. The fiber itself is constructed out of a stationary phase-like material which is selected for use based on the functionality of the substances present for extraction.

- ▶ Odor Biometrics

Spoofing

Spoofing entails the presentation of an artifact designed to imitate a legitimate biometric so as to defeat or circumvent a biometric system or process.

► Anti-spoofing
► Fraud Reduction, Overview

Spoofing Countermeasures

► Liveness Detection: Iris

Spoof-resistance

► Liveness Detection: Iris

Stand Off

The distance between a biometric capture device and the subject from whom the biometric is captured.

► Iris Device
► Iris on the Move

Statistical Models

► Deformable Models

Statistical Signal Processing

► Fusion, Decision-Level

Steganography

Hiding a secret message within a larger, typically unrelated cover message such that unintended recipients do not suspect the presence of the hidden message. In steganography, the cover message serves only as a mechanism to transport the secret message; it may be discarded or destroyed after the secret message has been extracted and communicated to the intended recipient. An example might include a spy posing as a tourist sending government secrets via images of tourist venues.

► Iris Digital Watermarking

Strain Gauge

Strain gauge is a device used to measure deformation (strain) of an object. It is a sensor whose resistance varies with applied force; it converts force, pressure, tension, weight, etc., into a change in electrical resistance which can then be measured.

► Digitizing Tablet

Stream of Speech

Stream of speech is the flow of sounds, words, and utterances produced by a human speaker.

► Speaker Recognition Standardization

Strength of Voice Evidence

The strength of the forensic evidence of voice is the result of the interpretation of the evidence, expressed in terms of the likelihood ratio of two alternative hypotheses: H_0 – the suspected speaker is the source of the questioned recording, H_1 – the speaker at the

origin of the questioned recording is not the suspected speaker.

► Voice, Forensic Evidence of

Stroke

Single movement of a pen during the signing process.

The number of strokes and the stroke order are very important features for on-line signature verification. Strokes are in general delimited by a change in the pen up/down status.

► Signature Recognition

Structural and Functional Anatomy

► Anatomy of Face

Structural Model

Describes the fundamental structures that interlink and constitute a body and specifies their constraints: e.g., torso and upper leg are linked by the hip, upper leg and lower leg are linked by the knee, and the knee constrains the degree of freedom of the lower leg. Therefore, it also defines the dependency of one body segment to another. Structural model is normally used together with a motion model to guide feature (of motion) extraction process. It may also consist of information such as length, thickness, area, or volume.

► Gait Recognition, Model-Based

Structural Risk

The error associated with the nature of the classifier. Structural risk minimization tunes the complexity of the classification function in order to optimize the generalization. The true risk combines structural risk with empirical risk, which is dependent on the training set.

► Support Vector Machine

Structure Tensor Field

The matrix field obtained by outer products of the gradients.

► Fingerprint Features

Structure-from-Motion

The process of recovering the three-dimensional structure of an object by the analysis of subsequent two-dimensional images of the object in motion. Although this term most commonly refers to the recovery of the structure of a rigid object (i.e., the distances between all points on the object remain constant), it still applies to piecewise rigid objects such as the human body where only some subsets of points remain rigid.

► Psychology of Gait and Action Recognition

Subject Interaction Time

► Operational Times

Super-Resolution

The techniques that form an enhanced-resolution image by fusing together multiple low-resolution

and/or learning from high-resolution training images are known as super-resolution. Super-resolution can be performed in either frequency or spatial domain.

► Face Sample Quality

Super-Resolution for Iris

► Iris Super-Resolution

Supervised

Supervised are the class labels of all data which are known. Algorithms are designed for separation of data of different classes by additionally using labeled information.

► Linear Dimension Reduction

Supervised Learning

Jong Kyoung Kim, Kye-Hyeon Kim, Seungjin Choi
Department of Computer Science, Pohang University of Science and Technology, Korea

Synonyms

Classification

Introduction

Supervised learning is to find a model for an unknown target function, using a given set of training examples – pairs of a data point and a target function value. The fundamental assumption of learning by examples is that similar data points tend to have similar function values. In this chapter, learning tasks are focused on classification, in which the aim is to assign each data point to one of a finite number of function values, called *class labels*. Then, the assumption can be restated that a group of similar data points form a meaningful pattern, which corresponds to the same class label.

Biometric recognition systems require two different classification problems of *verification* and *identification*, depending on the application context [1]. In the verification problems, we need to verify a claimed identity by comparing an input feature vector extracted from the biometric data with the corresponding template feature set. The verification problem can be restated as binary classification, in which the learned classifier takes the feature vector as an input and determines whether the claimed identity is true or false. The goal of identification problems, however, is to recognize a person's identity by searching templates corresponding to all the users enrolled in the biometric database. The identification problem may be stated as classification with $K+1$ classes $\{\mathcal{C}_1, \ldots, \mathcal{C}_K, \mathcal{C}_{K+1}\}$, where K is the number of the users enrolled in the system and the last class $\mathcal{C}_{K+1}$ indicates the unidentified user.

A variety of approaches for classification can be grouped into parametric and nonparametric, based on the assumptions about class-conditional densities [2]. If the class-conditional densities are specified with a functional form of distributions, which are characterized completely with a finite number of parameters, we have a parametric method for classification. In contrast, if we model the class-conditional densities without any assumptions regarding the functional form of the underlying distribution, non-parametric algorithms are derived. The methods that directly construct the decision boundaries without referring to the class-conditional densities are also considered as non-parametric approaches.

Parametric Approach

Suppose we have a D-dimensional real-valued feature vector $x = (x_1, \ldots x_D)^{\mathrm{T}}$. The goal of classification is to predict the class label for a new value of x given a training set of $\mathcal{D}$. According to Bayesian decision theory, an optimal decision boundary can be obtained from

posterior probabilities for classes. Using Bayes' theorem, the posterior probabilities $p(\mathcal{C}_k|x)$ can be written as

$$p(\mathcal{C}_k|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\mathcal{C}_k)p(\mathcal{C}_k)}{\sum_j p(\boldsymbol{x}|\mathcal{C}_j)p(\mathcal{C}_j)} = \frac{\exp(a_k)}{\sum_j \exp(a_j)}, \quad (1)$$

where $p(\boldsymbol{x}|\mathcal{C}_k)$ is the class-conditional density representing the probability of a feature vector x given class $\mathcal{C}_k$, $p(\mathcal{C}_k)$ is the prior probability of class $\mathcal{C}_k$, and a_k is defined by

$$a_k = \ln p(\boldsymbol{x}|\mathcal{C}_k)p(\mathcal{C}_k). \quad (2)$$

In parametric approaches to classification, we directly model the class-conditional density with a parametric form of probability distribution (e.g., multivariate Gaussian). Many parametric methods for classification have been proposed based on different assumptions for $p(\boldsymbol{x}|\mathcal{C}_k)$ [3, 4, 5] (see Table 1):

1. Linear discriminant analysis (LDA) uses Gaussian distributions $\mathcal{N}(\boldsymbol{\mu}_k, \Sigma_k)$ for the class-conditional densities with a common covariance matrix $\Sigma_k = \Sigma$.
2. Quadratic discriminant analysis (QDA) also uses Gaussian distributions and does not assume equal covariance matrices.
3. In a naive Bayes classifier, we assume that the distributions of the feature values $x_1, \ldots x_D$ are conditionally independent given the class label, $p(\boldsymbol{x}|\mathcal{C}_k) = \prod_{i=1}^{D} p(x_i|\mathcal{C}_k)$. In parametric approach, the one-dimensional densities $p(x_i|\mathcal{C}_k)$ are usually Gaussian for continuous features or multinomial for discrete features.

The parameters of each class-conditional densities can be estimated from a training set $\mathcal{D}$ with a maximum likelihood approach.

Supervised Learning. Table 1 Comparison among parametric methods for classification

Method	$p(\boldsymbol{x}\|\mathcal{C}_k)$	Number of parameters	Decision boundary
LDA	$\mathcal{N}(\boldsymbol{\mu}_k, \Sigma)$	$(K-1)$ $(D+1)$	Linear
QDA	$\mathcal{N}(\boldsymbol{\mu}_k, \Sigma_k)$	$(K-1)$ $(D(D+3)/2+1)$	Quadratic
Naive Bayes classifier with multinomial distributions	$\prod_{i=1}^{D} p(\boldsymbol{x}_i\|\mathcal{C}_k)$	$(K-1)$ $(D+1)$	Linear

Linear Discriminant Analysis

We assume that all class-conditional densities $p(\boldsymbol{x}|\mathcal{C}_k)$ are Gaussian with the same covariance matrix Σ. Then the density for class $\mathcal{C}_k$ is given by

$$p(\boldsymbol{x}|\mathcal{C}_k) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_k)^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_k)\right\}. \quad (3)$$

From (1), we have

$$a_k(\boldsymbol{x}) = \boldsymbol{w}_k^\top \boldsymbol{x} + \boldsymbol{w}_{k0}, \quad (4)$$

where

$$w_k = \Sigma^{-1}\boldsymbol{\mu}_k. \quad (5)$$

$$w_{k0} = -\frac{1}{2}\boldsymbol{\mu}_k^\top \Sigma^{-1}\boldsymbol{\mu}_k + \ln p(\mathcal{C}_k). \quad (6)$$

We see that the equal covariance matrices make $a_k(x)$ to be linear in x, and the resulting decision boundaries will also be linear. As a special case of LDA, the nearest-neighbor classifier can be obtained, when $\Sigma = \sigma^2 I$. If the prior probabilities $p(\mathcal{C}_k)$ are equal, we assign a feature vector x to the class $\mathcal{C}_k$ with the minimum Euclidean distance $||\boldsymbol{x} - \boldsymbol{\mu}_k||_2$, which is equivalent to the optimum decision rule based on the maximum posterior probability. Another extension of LDA could be obtained by allowing for mixtures of Gaussians for the class-conditional densities instead of the single Gaussian. Mixture discriminant analysis (MDA) [6] incorporates the Gaussian mixture distribution for the class-conditional densities to provide a richer class of density models than the single Gaussian. The class-conditional density for class $\mathcal{C}_k$ has the form of the Gaussian mixture model, $p(\boldsymbol{x}|\mathcal{C}_k) = \sum_{r=1}^{R} \pi_{kr}\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_k, \Sigma)$, where the mixing coefficients π_{kr} must satisfy $\pi_{kr} \geq 0$ together with $\sum_{r=1}^{R} \pi_{kr} = 1$. In this model, the same covariance matrix Σ is used within and between classes. The Gaussian mixture model allows for more complex decision boundaries although it does not guarantee the global optimum of maximum likelihood estimates.

Quadratic Discriminant Analysis

If the covariance matrices Σ_k are not assumed to be equal, then we get quadratic functions of $\boldsymbol{x}$ for $a_k(\boldsymbol{x})$

$$a_k(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_k)^\top \Sigma_k^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_k) - \frac{1}{2}\ln|\Sigma_k| + \ln p(\mathcal{C}_k). \quad (7)$$

In contrast to LDA, the decision boundaries of QDA are quadratic, which is resulted from the assumption on the different covariance matrices. From the added flexibility obtained from the quadratic decision boundaries, QDA often outperforms LDA when the size of training data is very large. However, when the size of the training set $\mathcal{D}$ is small compared to the dimension D of the feature space, the large number of parameters of QDA relative to LDA causes over-fitting or ill-posed estimation for the estimated covariance matrices. To solve this problem, various regularization or Bayesian techniques have been proposed to obtain more robust estimates:

1. Regularized discriminant analysis (RDA) [7, 8] employs the regularized form of covariance matrices by shrinking Σ_k of QDA towards the common covariance matrix Σ of LDA, that is, $\Sigma_k(\alpha) = \alpha\Sigma_k + (1-\alpha)\Sigma$ for $\alpha \in [0, 1]$. Additionally, the common covariance matrix Σ could be shrunk towards the scalar covariance, $\Sigma(\gamma) = \gamma\Sigma + (1-\gamma)\sigma^2 I$ for $\gamma \in [0, 1]$. The pair of parameters is selected by cross-validation based on the classification accuracy of the training set.
2. Leave-one-out covariance estimator (LOOC) [9] finds an optimal regularized covariance matrices by mixing four different covariance matrices of Σ_k, $\mathrm{diag}(\Sigma_k)$, Σ, and $\mathrm{diag}(\Sigma)$, where the mixing coefficients are determined by maximizing the average leave-one-out log likelihood of each class.
3. Bayesian QDA introduces prior distributions over the mean μ_k and the covariance matrices Σ_k [10], or over the Gaussian distributions themselves [11]. The expectations of the class-conditional densities are calculated analytically in terms of the parameters. The hyper-parameters of the prior distributions are chosen by cross-validation.

Naive Bayes Classifier

In the naive Bayes classifier, the conditional independence assumption makes the factorized class-conditional densities of the form

$$p(\boldsymbol{x}|\mathcal{C}_k) = \prod_{i=1}^{D} p(x_i|\mathcal{C}_k). \tag{8}$$

The component densities $p(\boldsymbol{x}_i|\mathcal{C}_k)$ can be modeled with various parametric and nonparametric distributions, including the following:

1. For continuous features, the component densities are chosen to be Gaussian. In this case, the naive Bayes classifier is equivalent to QDA with diagonal covariance matrices for each class.
2. For discrete features, multinomial distributions are used to model the component densities. The multinomial assumption makes $a_k(\boldsymbol{x})$ and the resulting decision boundaries to be linear in $\boldsymbol{x}$.
3. The component densities can be estimated using one-dimensional kernel density or histogram estimates for non-parametric approaches.

The naive Bayes model assumption is useful when the dimensionality D of the feature space is very high, making the direct density estimation in the full feature space unreliable. It is also attractive if the feature vector consists of heterogeneous features including continuous and discrete features.

Nonparametric Approaches

One major problem of parametric approaches is that the actual class-conditional density is not a linear nor a quadratic form in many real-world data. It causes the poor classification performance, since the actual distribution of data is different from a functional form we specified, regardless of parameters.

To solve this problem, one can increase the flexibility of the density model by adding more and more parameters, leading to a model with infinitely many number of parameters, called *nonparametric density estimation*. Otherwise, rather than modeling the whole distribution of a class, one can model only a *decision boundary* that separates one class from the others, since restricting the functional form of the boundary is a weaker assumption than restricting that of the whole distribution of data. Either using a nonparametric density model or modeling a decision boundary are called *nonparametric approaches*. In this article, the latter approach is only considered.

We define a function $a_k(\boldsymbol{x})$ as a relevancy score of $\boldsymbol{x}$ for $\mathcal{C}_k$, such that $a_k(\boldsymbol{x}) > 0$ if $\boldsymbol{x}$ is more likely to be assigned to $\mathcal{C}_k$, and $a_k(\boldsymbol{x}) < 0$ otherwise. Then, the surface $a_k(\boldsymbol{x}) = 0$ represents the decision boundary

Supervised Learning. Table 2 Comparison among non-parametric methods for classification

Method	$a_k(\boldsymbol{x})$	Number of parameters	Decision boundary
k-NN	$\lvert\{\boldsymbol{x}^{(i)} \in \mathcal{C}_k\}\rvert$	k	Nonlinear
ANNs	$f_k^{(L+1)}(\boldsymbol{x})$	$\sum_{\ell=0}^{L}(W_\ell+1)W_\ell+1$	Linear ($L=0$) or nonlinear ($L>0$)
SVMs	$\sum_{\alpha_{ki}>0}\alpha_{ki}y_{ki}k(\boldsymbol{x}_i,\boldsymbol{x})$	$O(K\,N)$	Linear ($k(\boldsymbol{x}_i,\mathrm{x})=\boldsymbol{x}_i^T\boldsymbol{x}$) or nonlinear (otherwise)

between $\mathcal{C}_k$ and the other classes, and a test point $\boldsymbol{x}$ is assigned to $\mathcal{C}_k$ if $k = \arg\max_k a_k(\boldsymbol{x})$, which is called *one-against-all*.

Many nonparametric methods have been derived from various models for $a_k(\mathrm{x})$. We introduce three representative methods [12, 13, 14] (see Table 2):

1. k-nearest neighbor algorithm (k-NN) chooses k data points in the training set, which are closest from x, then $a_k(\boldsymbol{x})$ is the number of those selected points belonging to $\mathcal{C}_k$.
2. Artificial neural networks (ANNs) represent $a_k(\boldsymbol{x})$ as a multilayered feed-forward network. The ℓth layer consists of W_ℓ nodes, where the jth node in the layer sends a (non)linear function value $f_j^{(\ell)}(\boldsymbol{x})$ as a signal to the nodes in the $(L+1)$th layer. Then, $a_k(\boldsymbol{x})$ is the signal of the kth node in the final layer, $f_k^{(L+1)}(\boldsymbol{x})$.
3. Support vector machines (SVMs) choose some "important" training points, called *support vectors*, then represent $a_k(\boldsymbol{x})$ as a linear combination of them. SVM is known to be the best supervised learning method for most real-world data.

k-Nearest Neighbor Algorithm

Given a set of data points $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N\}$ and a set of the corresponding labels $Y = \{y_1, y_2, y_N\}$, K-NN assigns a label for a test data point $\boldsymbol{x}$ by *majority voting*, that is to choose the most frequently occurred label in $\{y^{(1)}, y^{(2)}, \ldots y^{(k)}\}$, where $\boldsymbol{x}^{(i)}$ denotes the ith *nearest* point of $\boldsymbol{x}$ in X and $y^{(i)}$ is the label of $\boldsymbol{x}^{(i)}$. That is, we have

$$a_k(\boldsymbol{x}) = \lvert\{\boldsymbol{x}^{(i)} \in \mathcal{C}_k\}\rvert, \tag{9}$$

where $\lvert\cdot\rvert$ denotes the number of elements in a set. The decision boundary is not restricted to a specific functional form. It depends only on the local distribution of neighbors and the choice of k. Larger k makes the decision boundary more smooth.

k-NN is widely used in biometrics, especially for computer vision applications such as face recognition and pose estimation, where both the number of images N and dimension of data D are quite large. However, traditional k-NN takes $O(ND)$ time to compute distances between a test point x and all training points $\mathrm{x}_1, \ldots, \mathrm{x}_N$, which is too inefficient for practical use. Thus, extensive research has focused on fast approximations based on hashing, embedding or something [15].

Artificial Neural Networks

In ANNs, the signal of the jth node in the $(\ell+1)$th layer is determined by the signals from the ℓth layer:

$$f_j^{(\ell+1)}(\boldsymbol{x}) = g\left(\boldsymbol{w}_j^{(\ell)\top}\boldsymbol{f}^{(\ell)}(\boldsymbol{x}) + w_{j0}^{(\ell)}\right), \tag{10}$$

where $w_j^{\ell} = [w_{j1}^{(\ell)}, w_{j2}^{(\ell)}, \cdots, w_{jW_\ell}^{(\ell)}]^\top$ and $\boldsymbol{f}^{(\ell)}(\boldsymbol{x}) = [f_1^{(\ell)}(\boldsymbol{x}), f_2^{(\ell)}(\mathbf{x}), \cdots, f_{W_\ell}^{(\ell)}(\boldsymbol{x})]^\top$. The input layer, $\boldsymbol{f}^{(0)}(\boldsymbol{x})$, is simply $\boldsymbol{x}$. $g(\cdot)$ is a nonlinear, nondecreasing mapping, causing ANNs to yield a nonlinear decision boundary. There are two popular mappings: (1) sigmoid, $g(x) = 1/(1+exp\{-x\})$; (2) hyperbolic tangent, $g(x) = \tanh(x)$.

More nodes and layers increase the nonlinearity of decision boundary obtained by ANNs. However, it is difficult to train ANNs having a number of nodes and layers, since the model can easily fall into poor solutions, called *local minima*.

Radial basis function (RBF) networks [16] are another type of ANNs, having the form

$$a_k(\boldsymbol{x}) = \mathbf{w}_k^\top \Phi(\boldsymbol{x}) + \boldsymbol{w_{k0}}. \tag{11}$$

That is, RBF networks contain only one hidden layer, denoting by $\Phi(\boldsymbol{x}) = [\phi_1(\boldsymbol{x}), \phi_2(\boldsymbol{x}), \ldots, \phi_W(\boldsymbol{x})]$, and the network output is simply a linear combination of the hidden nodes. The main difference between RBF networks and ANNs with $L = 1$ is the mapping from the

input to the hidden. In RBF networks, each $\varphi_j(\cdot)$ is a nonlinear function similar to Gaussian density:

$$\phi_j(\boldsymbol{x}) = \exp\left\{-\beta_j||\boldsymbol{x} - \boldsymbol{c}_j||^2\right\}, \tag{12}$$

for some $\beta_j > 0$ and the center vector c_j. That is, each hidden node represents local region whose center is c_j, and its signal would be stronger if $\boldsymbol{x}$ and $\boldsymbol{c_j}$ are closer. In general, $\boldsymbol{c}_j$ is fixed to one of the training points and β_j is chosen by hand, thus the global optimum of $\boldsymbol{w}_k$ and w_{k0} can be simply found by least squares fitting.

Support Vector Machines

Similar to RBF networks, SVMs obtain a linear decision boundary in a transformed space: $a_k(\boldsymbol{x}) = \boldsymbol{w}_k^{\mathrm{T}}\Phi(x) + w_{k0}$, where $\Phi(\cdot)$ is an arbitrary mapping, either linear or nonlinear. The difference betwen SVMs and ANNs is the optimality of the decision boundary. In SVMs, the optimal decision boundary is such that the distance between the boundary and the closest point from that boundary, called the *margin*, is maximized:

$$\max_{\boldsymbol{w}_k, w_{k0}}\left[\min_i \frac{|a_k(\boldsymbol{x_i})|}{||\boldsymbol{w}_k||}\right]. \tag{13}$$

This optimization problem always converges to the global solution, the *maximum margin boundary*. Figure 1 shows the motivation for SVMs intuitively. One can expect that the generalization error of the maximum margin boundary is less than that of other boundaries. Theoretically, the generalization power of SVMs is guaranteed by *Vapnik–Chervonenkis theory* [17].

Training SVMs can be rewritten as the following convex optimization problem

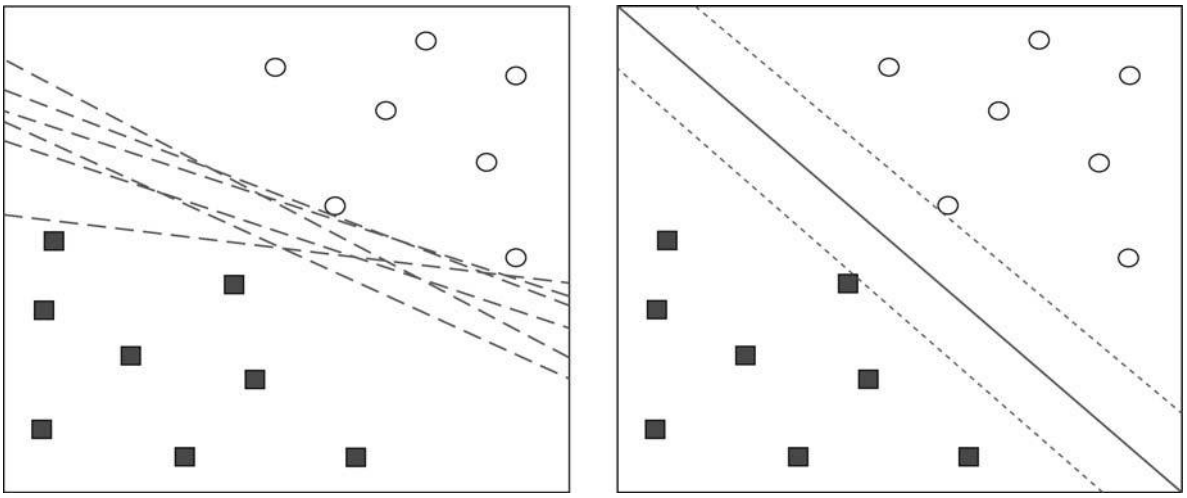
$$\min_{\boldsymbol{w}_k, w_{k0}} ||\boldsymbol{w}_k||, \quad \text{subject to} \quad y_{ki}a_k(\boldsymbol{x_i}) \geq \mathbf{1} \quad \text{for all } \mathbf{i}, \tag{14}$$

where $y_{ki} = 1$ if $\boldsymbol{x}_i \in \mathcal{C}_k$ and otherwise -1. At the optimum, $a_k(\boldsymbol{x})$ has the form

$$a_k(\boldsymbol{x}) = \sum_{i=1}^{n} \alpha_{ki} y_{ki} \Phi(\mathbf{x}_i)^{\top} \Phi(\mathbf{x}), \tag{15}$$

where $\alpha_{ki} \geq 0$ is a Lagrangian multiplier of the ith constraint, $y_{ki}a_k(\boldsymbol{x}_i) \geq 1$. If a data point x_i is exactly on the margin, i.e., $y_{ki}a_k(\boldsymbol{x}_i) = 1$, then $\boldsymbol{x}_i$ is called *support vector* and $\alpha_{ki} > 0$. Otherwise, $\alpha_{ki} = 0$ and $y_{ki}a_k(\boldsymbol{x}_i) > 1$. Hence, $a_k(\mathrm{x})$ only depends on the support vectors. To compute $\Phi(\boldsymbol{x}_i)^{\mathrm{T}}\Phi(\boldsymbol{x})$, we can introduce a function of the form $k(\boldsymbol{x}_i, \boldsymbol{x})$, representing the inner product in the feature space can be used, without computing the mapping $\Phi(\cdot)$ explicitly. Such a function is called *kernel function* [18]. There are two popular kernel functions: (1) *polynomial kernel*, $k(\boldsymbol{x}_i, \boldsymbol{x}) = (\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x} + c)^p$ for some c and $p > 0$; (2) *Gaussian kernel* (also called as *RBF kernel*), $k(\boldsymbol{x}_i, \boldsymbol{x}) = \exp\{-\frac{1}{2\sigma^2}||\boldsymbol{x}_i - \boldsymbol{x}||^2\}$ for some $\sigma > 0$.

Various algorithms and implementations have been developed to train SVMs efficiently. Two most popular softwares are LIBSVM [19] and SVM^{light} [20], both



Supervised Learning. Figure 1 (*Left*) Possible solutions obtained by neural networks. (*Right*) SVMs give one global solution, the maximum margin boundary.

implement several techniques such as working set selection, shrinking heuristics, and LRU caching to speed up optimization, and provide various kernel functions with choosing appropriate parameters of those functions automatically (automatic model selection). Two recent extensions of SVM^{light} – SVM^{struct} for structured data, SVM^{perf} for training with more than hundred-thousands of data points – are also popular in biometrics.

Related Entries

- ▶ Classifier Design
- ▶ Machine-Learning
- ▶ Probability Distribution

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 4–20 (2004)
2. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: A review. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1), 4–37 (2000)
3. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley, New York (2001)
4. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)
5. Hastie, T., Tibsjirani, R., Friedman, J.: The Elements of Statistical Learning. Springer, New York (2001)
6. Hastie, T., Tibshirani, R.: Discriminant analysis by Gaussian mixtures. J. R. Stat. Soc. Ser. B **58**, 158–176 (1996)
7. Friedman, J.H.: Regularized discriminant analysis. J. Am. Stat. Assoc. **84**, 165–175 (1989)
8. Ye, J., Wang, T.: Regularized discriminant analysis for high dimensional, low sample size data. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Philadelphia, PA (2006)
9. Hoffbeck, J.P., Landgrebe, D.A.: Covariance matrix estimation and classification with limited training data. IEEE Trans. Pattern Anal. Mach. Intell. **18(7)**, 763–767 (1996)
10. Geisser, S.: Predictive Inference: An Introduction. Chapman & Hall, New York (1993)
11. Srivastava, S., Gupta, M.R., Frigyik, B.A.: Bayesian quadratic discriminant analysis. J. Mach. Learn. Res. **8**, 1277–1305 (2007)
12. Cover, T., Hart, P.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theory **IT–13**, 21–27 (1967)
13. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by backpropagating errors. Nature **323**, 533–536 (1986)
14. Boser, B.E., Guyon, I., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop of Computational Learning Theory, pp. 144–152 (1992)
15. Shakhnarovich, G., Darrell, T., Indyk, P.: Nearest-Neighbor Methods in Learning and Vision: Theory and Practice. MIT Press, Cambridge, MA (2006)
16. Moody, J., Darken, C.J.: Fast learning in networks of locally tuned processing units. Neural Comput. **1**, 281–294 (1989)
17. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer-Verlag, New York (1995)
18. Schölkopf, B., Smola, A.J.: Learning with Kernels. MIT Press, Cambridge, MA (2002)
19. Chang, C.C., Lin, C.J.: LIBSVM – A Library for Support Vector Machines, http://www.csie.ntu.edu.tw/~cjlin/libsvm (2000)
20. Joachims, T.: SVM^{light}, http://svmlight.joachims.org (2004)

Supervisor

A generic term for a method or a system that is able to output an aggregated opinion.

- ▶ Multiple Experts

Supervisor Opinion

The output of the supervisor which can be a strict score (0 or 1) or a graded score ($\in [0, 1]$) representing the belief of the supervisor on an identity claim by aggregating expert opinions.

- ▶ Multiple Experts

Support Vector Machine

Mathias M. Adankon, Mohamed Cheriet
University of Quebec ETS, Montreal, Canada

Synonyms

SVM; Margin classifier; Maximum margin classifier; Optimal hyperplane

Definition

Support vector machines (SVMs) are particular linear ▶ classifiers which are based on the margin maximization principle. They perform ▶ structural risk minimization, which improves the complexity of the classifier with the aim of achieving excellent ▶ generalization performance. The SVM accomplishes the classification task by constructing, in a higher dimensional space, the hyperplane that optimally separates the data into two categories.

Introduction

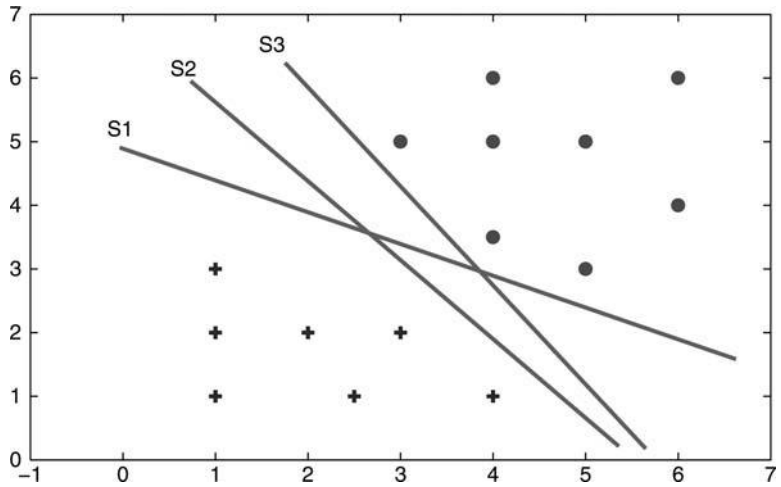
Considering a two-category classification problem, a linear classifier separates the space, with a hyperplane, into two regions, each of which is also called a class. Before the creation of SVMs, the popular algorithm for determining the parameters of a linear classifier was a single-neuron perceptron. The perceptron algorithm uses an updating rule to generate a separating surface for a two-class problem. The procedure is guaranteed to converge when the ▶ training data are linearly separable, however there exists an infinite number of hyperplanes that correctly classify these data (see Fig. 1).

The idea behind the SVM is to select the hyperplane that provides the best generalization capacity. Then, the SVM algorithm attempts to find the maximum margin between the two data categories and then determines the hyperplane that is in middle of the maximum margin. Thus, the points nearest the decision boundary are located at the same distance from the optimal hyperplane. In machine learning theory, it is demonstrated that the margin maximization principle provides the SVM with a good generalization capacity, because it minimizes the structural risk related to the complexity of the SVM [1].

SVM Formulation

Let consider a dataset $\{(x_1, y_1), \ldots, (x_\ell, y_\ell)\}$ with $x_i \in \mathcal{R}^d$ and $y_i \in \{-1, 1\}$. SVM training attempts to find the parameters w and b of the linear decision function $f(x) = w.x + b$ defining the optimal hyperplane. The points near the decision boundary define the margin. Considering two points x_1, x_2 on opposite sides of the margin with $f(x_1) = 1$ and $f(x_2) = -1$, the margin equals $[f(x_1) - f(x_2)]/\|w\| = 2/\|w\|$. Thus, maximizing the margin is equivalent to minimizing $\|w\|/2$ or $\|w\|^2/2$. Then, to find the optimal hyperplane, the SVM solves the following optimization problem:

$$\min_{w,b} \frac{1}{2} w'w \quad \text{s.t } y_i(w'.x_i + b) \geq 1 \ \forall i = 1, \ldots, \ell \tag{1}$$



Support Vector Machine. Figure 1 Linear classifier: In this case, there exists an infinite number of solutions. Which is the best?

The transformation of this optimization problem into its corresponding dual problem gives the following quadratic problem:

$$\max_{\propto} \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i y_i (x_i . x_j)$$
$$\text{s.t} \sum_{i=1}^{\ell} y_i \alpha_i = 0;\ \alpha \geq 0\ \forall i = 1, \ldots, \ell \tag{2}$$

Where W' denotes the transpose of W.

The solution of the previous problem gives the parameter $w = \sum_{i=1}^{\ell} y_i \alpha_i x_i$ of the optimal hyperplane. Thus, the decision function becomes $f(x) = \sum_{i=1}^{\ell} \alpha_i y_i (x_i . x) + b$ in dual space. Note that the value of the bias b does not appear in the dual problem. Using the constraints of the primal problem, the bias is given by $b = -1/2[max_{y=-1}(w.x_i) + min_{y=1}(w.x_i)]$. It is demonstrated with the Karush-Kuhn-Tucker conditions that only the examples x_i that satisfy $y_i(w.\ x_i + b) = 1$ are the corresponding α_i non-zero. These examples are called *support vectors* (see Fig. 2).

SVM in Practice

In real-world problems, the data are not linearly separable, and so a more sophisticated SVM is used to solve them. First, the slack variable is introduced in order to relax the margin (this is called a soft margin optimization). Second, the kernel trick is used to produce nonlinear boundaries [2]. The idea behind kernels is to map training data nonlinearly into a higher-dimensional feature space via a mapping function Φ and to construct a separating hyperplane which maximizes the margin (see Fig. 3). The construction of the linear decision surface in this feature space only requires the evaluation of dot products $\phi(x_i).\phi(x_j) = k(x_i, x_j)$, where the application $k : \mathcal{R}^d \times \mathcal{R}^d \rightarrow \mathcal{R}$ is called the kernel function [3, 4].

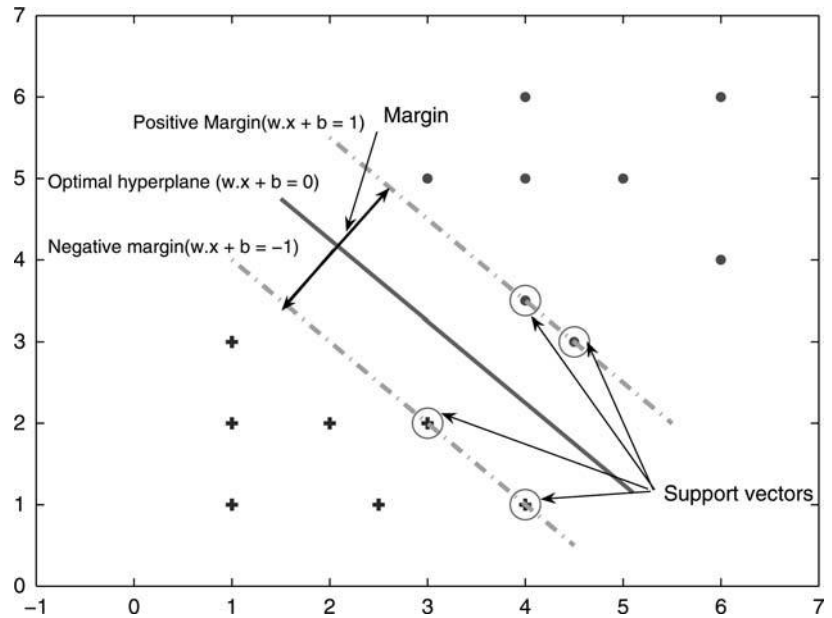
The decision function given by an SVM is:

$$y(x) = sign[w'\phi(x) + b], \tag{3}$$

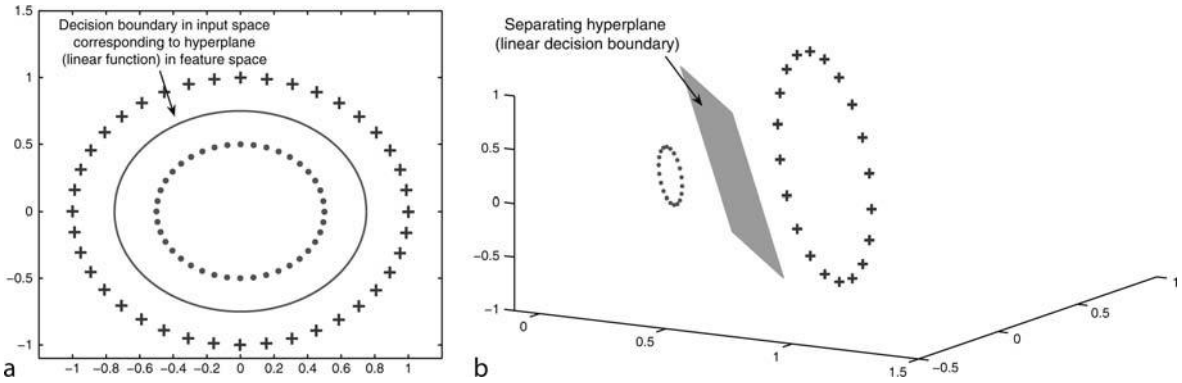
where w and b are found by resolving the following optimization problem that expresses the maximization of the margin $2/||w||$ and the minimization of training error:

$$\min_{w,b,\xi} \frac{1}{2} w'w + C \sum_{i=1}^{\ell} \xi_i\ (\text{L1} - \text{SVM}) \quad \text{or}$$
$$\min_{w,b,\xi} \frac{1}{2} w'w + C \sum_{i=1}^{\ell} \xi_i^2\ (\text{L2} - \text{SVM}) \tag{4}$$

$$\text{subject to} : y_i[w'\phi(x_i) + b] \geq 1 - \xi_i\ \forall i = 1, \ldots, \ell \tag{5}$$



Support Vector Machine. Figure 2 SVM principle: illustration of the unique and optimal hyperplane in a two-dimensional input space based on margin maximization.



Support Vector Machine. **Figure 3** Illustration of the kernel trick: The data are mapped into a higher-dimensional feature space, where a separating hyperplane is constructed using the margin maximization principle. The hyperplane is computed using the kernel function without the explicit expression of the mapping function. **(a)** Nonlinearly separable data in the input space. **(b)** Data in the higher-dimensional feature space.

$$\xi_i \geq 0 \ \forall i = 1, \ldots, \ell. \tag{6}$$

By applying the Lagrangian differentiation theorem to the corresponding dual problem, the following decision function is obtained:

$$y(x) = sign[\sum_{i=1}^{\ell} \alpha_i y_i k(x_i, x) + b], \tag{7}$$

with α solution of the dual problem.

The dual problem for the L1-SVM is the following quadratic optimization problem:

$$maximize \ : \ W(\alpha) = \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i y_j k(x_i, x_j) \tag{8}$$

$$subject\ to: \sum_{i=1}^{\ell} \alpha_i y_i = 0 \ and \ 0 \leq \alpha_i \leq C, i = 1, \ldots, \ell. \tag{9}$$

Using the L2-SVM, the dual problem becomes :

$$maximize \ : \ W(\alpha) = \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i.j=1}^{\ell} \alpha_i \alpha_j y_i y_j \left(k(x_i, x_j) + \frac{1}{2C} \delta_{ij} \right) \tag{10}$$

$$subject\ to: \sum_{i=1}^{\ell} \alpha_i y_i = 0 \ and \ 0 \leq \alpha_i, i = 1, \ldots \ell. \tag{11}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

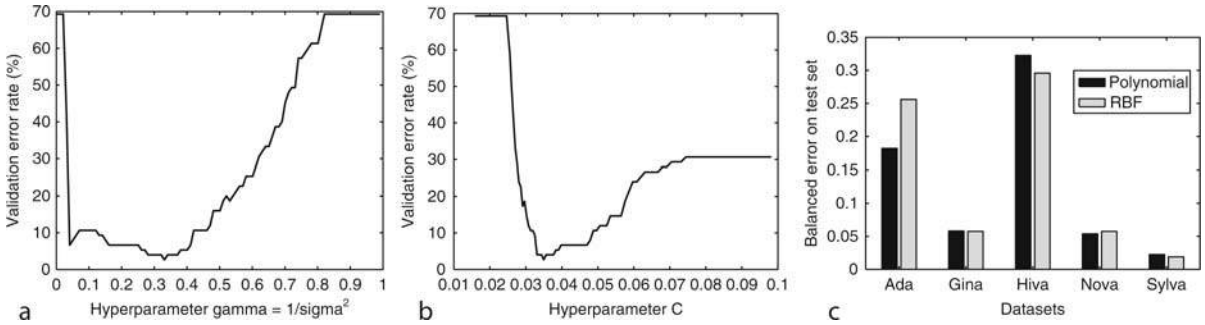
Support Vector Machine. **Table 1** Common kernel used with SVM

Gaussian (RBF)	$k(x, y) = exp(-\|\|x - y\|\|^2/\sigma^2)$
Polynomial	$k(x, y) = (ax.y + b)^n$
Laplacian	$(x, y) = exp(-a\|\|x - y\|\| + b)$
Multi-quadratic	$k(x, y) = (a\|\|x - y\|\| + b)^{1/2}$
Inverse multi-quadratic	$k(x, y) = (a\|\|x - y\|\| + b)^{-1/2}$
KMOD	$k(x, y) = a\left[exp\left(\frac{\gamma^2}{\|\|x-y\|\|^2+\sigma^2}\right) - 1\right]$

In practice, the L1-SVM is used most of the time, and its popular implementation developed by Joachims [5] is very fast and scales to large datasets. This implementation, called *SVMlight*, is available at svmlight.joachims.org.

SVM Model Selection

To achieve good SVM performance, optimum values for the kernel parameters and for the hyperparameter C much be chosen. The latter is a regularization parameter controlling the trade-off between the training error minimization and the margin maximization. The kernel parameters define the kernel function used to map data into a higher-dimensional feature space (see Table 1). Like kernel functions, there are the Gaussian kernel $k(x_i, x_j) = exp(-\|x_i - x_j\|^2/\sigma^2)$ with parameter σ and



Support Vector Machine. Figure 4 **(a)** and **(b)** show the impact of SVM hyperparameters on classifier generalization, while **(c)** illustrates the influence of the choice of kernel function.

the polynomial kernel $k(x_i, x_j) = (ax_i' x_j + b)^d$ with parameters a, b and d. The task of selecting the hyperparameters that yield the best performance of the machine is called model selection [6, 7, 8, 9].

As an illustration, Fig. 4a shows the variation of the error rate on a validation set versus the variation of the Gaussian kernel with a fixed value of C and Fig. 4b shows the variation of the error rate on the validation set versus the variation of the hyperparameter C with a fixed value of the RBF kernel parameter. In each case, the binary problem described by the "Thyroid" data taken from the UCI benchmark is resolved. Clearly, the best performance is achieved with an optimum choice of the kernel parameter and of C.

With the SVM, as with other kernel classifiers, the choice of kernel corresponds to choosing a function space for learning. The kernel determines the functional form of all possible solutions. Thus, the choice of kernel is very important in the construction of a good machine. So, in order to obtain a good performance from the SVM classifier, one first need to design or choose a type of kernel, and then optimize the SVM's hyperparameters to improve the generalization capacity of the classifier. Figure 4c illustrates the influence of the kernel choice, where the RBF and the polynomial kernels are compared on datasets taken from the challenge website on model selection and prediction organized by Isabelle Guyon.

Resolution of Multiclass Problems with the SVM

The SVM is formulated for the binary classification problem. However, there are some techniques used to combine several binary SVMs in order to build a system for the multiclass problem (e.g., a 10-class digit recognition problem). Two popular methods are presented here:

OneVersustheRest: The idea of one versus the rest is to construct as many SVMs as there are classes, where each SVM is trained to separate one class from the rest. Thus, for a c-class problem, c SVMs are built and combined to perform multiclass classification according to the maximal output. The *ith* SVM is trained with all the examples in the *ith* class with positive labels, and all the other examples with negative examples. This is also known as the *One-Against-All* method.

Pairwise(*orOne* − *Against* − *One*): The idea of pairwise is to construct $c(c-1)2$ SVMs for a c-class problem, each SVM being trained for every possible pair of classes. A common way to make a decision with the pairwise method is by voting. A rule for discriminating between every pair of classes is constructed, and the class with the largest vote is selected.

SVM Variants

The least squares SVM (LS-SVM) is a variant of the standard SVM, and constitutes the response to the following question: *How much can the SVM formulation be simplified without losing any of its advantages?* Suykens and Vandewalle [10] proposed the LS-SVM where the training algorithm solves a convex problem like the SVM. In addition, the training algorithm of the LS-SVM is simplified, since a linear problem is resolved instead of a quadratic problem in the SVM case.

The Transductive SVM (TSVM) is an interesting version of the SVM, which uses transductive inference. In this case, the TSVM attempts to find the hyperplane and the labels of the test data that maximize the margin with minimum error. Thus, the label of the test data is obtained in one step. Vapnik [1] proposed this formulation to reinforce the classifier on the test set by adding the minimization of the error on the test set during the training process. This formulation has been used elsewhere recently for training semi-supervised SVMs.

Applications

The SVM is a powerful classifier which has been used successfully in many pattern recognition problems, and it has also been shown to perform well in biometrics recognition applications. For example, in [11], an iris recognition system for human identification has been proposed, in which the extracted iris features are fed into an SVM for classification. The experimental results show that the performance of the SVM as a classifier is far better than the performance of a classifier based on the artificial neural network. In another example, Yao et al. [12], in a fingerprint classification application, used recursive neural networks to extract a set of distributed features of the fingerprint which can be integrated into the SVM. Many other SVM applications, like handwriting recognition [8, 13], can be found at www.clopinet.com/isabelle/Projects/SVM/applist.html.

Related Entries

- ▶ Classifier
- ▶ Generalization
- ▶ Structural Risk
- ▶ Training

References

1. Vapnik, V.N.: Statistical learning theory. Wiley, New York (1998)
2. Boser, B.M.E., Guyon, I., Vapnik, V.: A training algorithm for optimal margin classifiers. In: Proceedings of Fifth Annual Workshop on Computational Learing Theory, pp. 144–152 (1992)
3. Scholkopf, B., Smola, A.J.: Learning with Kernels. MIT Press, Cambridge, MA (2002)
4. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines. Cambridge University Press (2000)
5. Joachims, T.: Making large-scale support vector machine learning practical. In: Scholkopf, Burges, Smola (eds.) Advances in Kernel Methods: Support Vector Machines. MIT Press, Cambridge, MA (1998)
6. Chapelle, O., Vapnik, V.: Model selection for support vector machines. Advances in Neural Information Processing Systems (1999)
7. Ayat, N.E., Cheriet, M., Suen, C.Y.: Automatic Model Selection for the Optimization of the SVM kernels. Pattern Recognit. **38**(10), 1733–1745 (2005)
8. Adankon, M.M., Cheriet, M.: Optimizing Resources in Model Selection for Support Vector Machines. Pattern Recognit. **40**(3), 953–963 (2007)
9. Adankon, M.M., Cheriet, M.: New formulation of svm for model selection. In: IEEE International Joint Conference in Neural Networks 2006, pp. 3566–3573. Vancouver, BC (2006)
10. Suykens, J.A.K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J.: Least Squares Support Vector Machines. World Scientific, Singapore (2002)
11. Roy, K., Bhattacharya, P.: Iris recognition using support vector machine. In: APR International Conference on Biometric Authentication (ICBA), Hong Kong, January 2006. Springer Lecture Note Series in Computer Science (LNCS), pp. (3882) 486–492 (2006)
12. Yao, Y., Marcialis, G.L., Pontil, M., Frasconi, P., Rolib, F.: Combining flat and structured representations for fingerprint classification with recursive neural networks and support vector machines. Pattern Recognit. **36**(2), 397–406 (2003)
13. Matic, N., Guyon, I., Denker, J., Vapnik, V.: Writer adaptation for on-line handwritten character recognition. In: IEEE Second International Conference on Pattern Recognition and Document Analysis, pp. 187–191. Tsukuba, Japan (1993)

Surface Curvature

Measurements of the curvature of a surface are commonly used in 3D biometrics. The *normal curvature* on a point p on the surface is defined as the curvature of the curve that is formed by the intersection of the surface with the plane containing the normal vector and one of the tangent vectors at p. Thus the normal curvature is a function of the tangent vector direction. The minimum and maximum values of this function are the *principal curvatures* $k1$ and $k2$ of the surface

at p. Other measures of surface curvature are the *Gaussian curvature* defined as the product of principal curvatures, the *mean curvature* defined as the average of principal curvatures and the *shape index* given by

$$SI = \frac{2}{\pi}\frac{k_2 + k_1}{k_2 - k_1}$$

Computation of surface curvature on discrete surfaces such as those captured with 3D scanners is usually accomplished by locally fitting low order surface patches (e.g. biquadratic surfaces, splines) over each point. Then the above curvature features may be computed analytically.

► Finger Geometry, 3D

Surface Matching

3D biometrics work by computing the similarity between 3D surfaces of objects belonging to the same class. The majority of the techniques used measure the similarity among homologous salient geometric features on the surfaces (e.g. based on curvature). The localization of these features is usually based on prior knowledge of the surface class (e.g. face, hand) and thus, specialized feature detectors may be used. The geometric attributes extracted are selected so that they are invariant to transformations such as rotation, translation and scaling. In the case that knowledge-based feature detection is difficult, a correspondence among the surfaces may be established by randomly selecting points on the two surfaces and then trying to find pairs of points with similar geometric attributes. Several such techniques have been developed for rigid surface matching (e.g. Spin Images) which may be extended for matching non-rigid or articulated surfaces. Another technique for establishing correspondences is fitting a parameterized deformable model to the points of each surface. Since the fitted models are deformations of the same surface, correspondence is automatically determined. Creation of such deformable models requires however a large number of annotated training data.

► Finger Geometry, 3D

Surveillance

Rama Chellappa, Aswin C. Sankaranarayanan
University of Maryland, College Park, MD, USA

Synonyms

Monitoring; Surveillance

Definition

Surveillance refers to monitoring of a scene along with analysis of behavior of the people and vehicles for the purpose of maintaining security or keeping a watch over an area. Typically, traditional surveillance involves monitoring of a scene using one or more close circuit television (CCTV) cameras with personnel watching and making decisions based on video feeds obtained from the ► cameras. There is a growing need towards building systems that are completely automated or operate with minimal human supervision.

Biometric acquisition and processing is by far the most important component of any *automated* surveillance system. There are many challenges and variates that show up in acquisition of biometrics for robust verification. Further, in surveillance, behavioral biometrics is also of potential use in many scenarios. Using the patterns observed in a scene (such as faces, speech, behavior), the system decides on a set of actions to perform. These actions could involve access control (allowing/denying access to facilities), alerting the presence of intruders/abandoned luggage and a host of other security related tasks.

Introduction

Surveillance refers to monitoring a scene using sensors for the purposes of enhanced security. Surveillance systems are becoming ubiquitous, especially in urban areas with growing deployment of cameras and CCTV for providing security in public areas such as banks, shopping malls, etc. It is estimated that UK alone has more than four million CCTV cameras. Surveillance technologies are also becoming common for

other applications [1] such as traffic monitoring, wherein it is mainly used for detecting violations and monitoring traffic. Typically, video cameras are finding use for detecting congestion, accidents, and in adaptive switching of traffic lights. Other typical surveillance tasks include portal control, monitoring shop lifting, and suspect tracking as well as post-event analysis [2].

A traditional surveillance system involves little automation. Most surveillance systems have a set of cameras monitoring a scene of interest. Data collected from these sensors are used for two purposes.

1. Real time monitoring of the scene by human personnel.
2. Archiving of data for retrieval in the future.

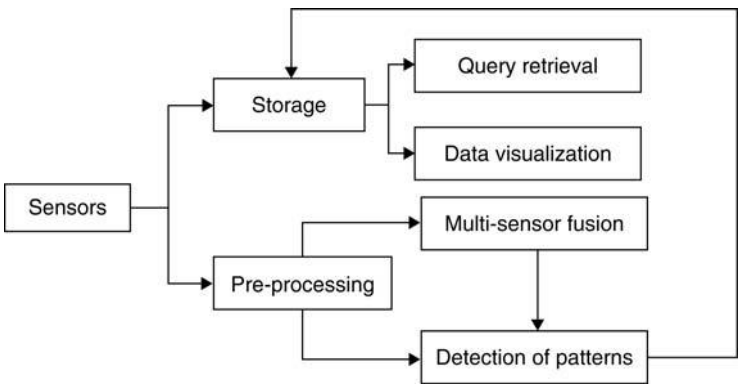
In most cases, the archived data is only retrieved after an incident has occurred.

This, however is changing with introduction of many commercial surveillance technologies that introduce more automation thereby alleviating the need or reducing the involvement of humans in the decision making process [3]. Simultaneously, the focus has also been in visualization tools for better depiction of data collected by the sensors and in fast retrieval of archived data for quick forensic analysis. Surveillance systems that can detect elementary events in the video streams acquired by several cameras are commercially available today. A very general surveillance system is schematically shown in Fig. 1.

Biometrics form a critical component in all (semi-)automated surveillance systems, given the obvious need to acquire, validate, and process biometrics in various surveillance tasks. Such tasks include:

1. *Verification.* Validating a person's identity is useful in access control. Typically, verification can be done in a controlled manner, and can use active biometrics such as iris, face (controlled acquisition), speech, finger/hand prints. The system is expected to use the biometrics to confirm if the person is truly whom he/she claims to be.
2. *Recognition.* Recognition of identity shows up in tasks of intruder detection and screening, which finds use in a wide host of scenarios from scene monitoring to home surveillance. This involves cross-checking the acquired biometrics across a list to obtain a match. Typically, for such tasks, passive acquisition methods are preferred making face and gait biometrics useful for this task.
3. *Abnormality detection.* Behavioral biometrics find use in surveillance of public areas, such as airports and malls, where the abnormal/suspicious behavior exhibited by a single or group of individual forms is the biometric of interest.

Biometrics finds application across a wide range of surveillance tasks. We next discuss the variates and trade-offs involved in using biometrics application for surveillance.



Surveillance. **Figure 1** Inputs from sensors are typically stored on capture. The relevant information is searched and retrieved only after incidents. However, in more automated systems the inputs are pre-processed for events. The system monitors these certain patterns to occur which initiates the appropriate action. When multiple sensors are present, for additional robustness, data across sensors might be fused.

Biometrics and Surveillance

The choice of biometric to be used in a particular task depend on the match between the acquisition and processing capability of the biometric to the requirements of the task. Such characteristics include the discriminative power of the biometric, ease of acquisition, the permanence of the biometric, and miscellaneous considerations such as acceptability of its use and ▶ privacy concerns [4, 5]. Towards this end, we discuss some of the important variates that need be considered in biometric surveillance.

1. *Cooperative acquisition.* Ease of acquisition is probably the most important consideration for use of a particular biometric. Consider the task of home surveillance, where the system tries to detect intruders by comparing the acquired biometric signature to a database of individuals. It is not possible in such a task to use iris as a biometric, because acquisition of iris pattern requires cooperation of the subject. Similarly, for the same task, it is also unreasonable to use controlled face recognition (with known pose and illumination) as a possible biometric for similar reasons.

Using the cooperation of subject as a basis, allows us to classify biometrics into two kinds: *cooperative* and *non-cooperative*. Fingerprints, hand prints, speech (controlled), face (controlled), iris, ear, and DNA are biometrics that need the active cooperation of the subject for acquisition. These biometrics, given the cooperative nature of acquisition, can be collected reliably under a controlled setup. Such controls could be a known sentence for speech, a known pose and favorable illumination for face. Further, the subject could be asked for multiple samples of the same biometric for increased robustness to acquisition noise and errors. In return, it is expected that the biometric performs at increased reliability with lower false alarms and lower mis-detections. However, the cooperative nature of acquisition makes these biometrics unusable for a variety of operating tasks. None the less, such biometrics are extremely useful for a wide range of tasks, such as secure access control, and for controlled verification tasks such as those related to passports and other identification related documents.

In contrast, acquisition of the biometric without the cooperation of the subject(s) is necessary for surveillance of regions with partially or completely unrestricted access, wherein the sheer number of subjects involved does not merit the use of active acquisition. Non-cooperative biometrics are also useful in surveillance scenarios requiring the use of behavioral biometrics, as with behavioral biometrics the use of active acquisition methods might inherently affect the very behavior that we want to detect. Face and gait are probably the best examples of such biometrics.

2. *Inherent capability of discrimination.* Each biometric depending on its inherent variations across subjects, and intra-variations for each individual has limitations on the size of the dataset it can be used before its operating characteristics (false alarm and mis-detection rate) go below acceptable limits. DNA, iris, and fingerprint provide robust discrimination even when the number of individuals in the database are in tens of thousands. Face (under controlled acquisition) can robustly recognize with low false alarms and mis-detections upto datasets containing many hundreds of individuals. However, performance of face as a biometric steeply degrades with uncontrolled pose, illumination, and other effects such as aging, disguise, and emotions. Gait, as a biometric provides similar performance capabilities as that of face under uncontrolled acquisition. However, as stated above, both face and gait can be captured without the cooperation of the subject, which makes them invaluable for certain applications. However, their use also critically depends on the size of the database that is used.

3. *Range of operation.* Another criterion that becomes important in practical deployment of systems using biometrics is the range at which acquisition can be performed. Gait, as an example, works with the human silhouette as the basic building block, and can be reliably captured at ranges upto a 100 m (assuming a common deployment scenario). In contrast, fingerprint needs contact between the subject and the sensor. Similarly, iris requires the subject to be at much closer proximity than what is required for face.

4. *Miscellaneous considerations.* There exist a host of other considerations that decide the suitability of a biometric to a particular surveillance application. These include the permanence of the biometric, security considerations such as the ease of imitating or tampering, and privacy considerations in its acquisition and use [4, 5]. For example, the permanence of face as a biometric depends on the degradation of its discriminating capabilities as the subject ages [6, 7].

Similarly, the issue of wear of the fingerprints with use becomes an issue for consideration. Finally, privacy considerations play an important role in the acceptability of the biometrics' use in commercial systems.

Behavioral Biometrics in Surveillance

Behavioral biometrics are very important for surveillance, especially towards identifying critical events before or as they happen. In general, the visual modality (cameras) is most useful for capturing behavioral information, although there has been some preliminary work on using motion sensor for similar tasks. In the presence of a camera, the processing of data to obtain such biometrics falls under the category of event detection. In the context of surveillance systems, these can be broadly divided into those that model actions of single objects and those that handle multi-object interactions. In the case of single-objects, an understanding of the activity (behavior) being performed is of immense interest. Typically, the object is described in terms of a feature-vector [8] whose representation is suitable to identify the activities while marginalizing nuisance parameters such as the identity of the object or view and illumination. Stochastic models such as the Hidden Markov models and Linear Dynamical Systems have been shown to be efficient in modeling activities. In these, the temporal dynamics of the activity are captured using state-space models, which form a generative model for the activity. Given a test activity, it is possible to evaluate the likelihood of the test sequence arising from the learnt model.

Capturing the behavioral patterns exhibited by multiple actors is of immense importance in many surveillance scenarios. Examples of such interactions include an individual exiting a building and driving a car, or an individual casing vehicles. A lot of other scenarios, such as abandoned vehicles, dropped objects fit under this category. Such interactions can be modeled using context-free grammars [9, 10] (Fig. 2). Detection and tracking data are typically parsed by the rules describing the grammar and a likelihood of the particular sequence of tracking information conforming to the grammar is estimated. Other approaches rely on motion analysis of humans accompanying the abandoned objects.

The challenges towards the use of behavioral biometrics in surveillance tasks are in making algorithms robust to variations in pose, illumination, and identity. There is also the need to bridge the gap between the tools for representation and processing used for identifying biometrics exhibited by individuals and those by groups of people. In this context, motion sensors [11] provide an alternate way for capturing behavioral signatures of groups. Motion sensors register time-instants when the sensor observes motion in its range. While this information is very sparse, without any ability to recognize people or disambiguate between multiple targets, a dense deployment of motion sensors along with cameras can be very powerful.

Conclusion

In summary, biometrics are an important component of automated surveillance, and help in the tasks of recognition and verification of a target's identity. Such tasks find application in a wide range of surveillance applications. The use of a particular biometric for a surveillance application depends critically on the match between the properties of the biometrics and the needs of the application. In particular, attributes

Surveillance. **Figure 2** Example frames from a detected casing incident in a parking lot. The algorithm described in [10] was used to detect the casing incident.

such as ease of acquisition, range of acquisition and discriminating power form important considerations towards the choice of biometric used. In surveillance, behavioral biometrics are useful in identifying suspicious behavior, and finds use in a range of scene monitoring applications.

Related Entries

- ▶ Border Control
- ▶ Law Enforcement
- ▶ Physical Access Control
- ▶ Face Recognition, Video Based

References

1. Remagnino, P., Jones, G.A., Paragios, N., Regazzoni, C.S.: Video-Based Surveillance Systems: Computer Vision and Distributed Processing. Kluwer, Dordecht (2001)
2. Zhao, W., Chellappa, R., Phillips, P., Rosenfeld, A.: Face recognition: a literature survey. ACM Comput. Surv. **35**, 399–458 (2003)
3. Shu, C., Hampapur, A., Lu, M., Brown, L., Connell, J., Senior, A., Tian, Y.: Ibm smart surveillance system (s3): a open and extensible framework for event based surveillance. In: IEEE Conference on Advanced Video and Signal Based Surveillance. pp. 318–323 (2005)
4. Prabhakar, S., Pankanti, S., Jain, A.: Biometric recognition: security and privacy concerns. Security & Privacy Magazine, IEEE **1**, 33–42 (2003)
5. Liu, S., Silverman, M.: A practical guide to biometric security technology. IT Professional **3**, 27–32 (2001)
6. Ramanathan, N., Chellappa, R.: Face verification across age progression. Comput. Vision Pattern Recognit., 2005. CVPR 2005. IEEE Computer Society Conference on **2** (2005)
7. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.: A study of face recognition as people age. Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on pp. 1–8 (2007)
8. Veeraraghavan, A., Roy-Chowdhury, A.K., Chellappa, R.: Matching shape sequences in video with applications in human movement analysis. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1896–1909 (2005)
9. Moore, D., Essa, I.: Recognizing multitasked activities from video using stochastic context-free grammar. Workshop on Models versus Exemplars in Computer Vision (2001)
10. Joo, S., Chellappa, R.: Recognition of multi-object events using attribute grammars. IEEE International Conference on Image Processing pp. 2897–2900 (2006)
11. Wren, C., Ivanov, Y., Leigh, D., Westhues, J.: The MERL Motion Detector Dataset: 2007 Workshop on Massive Datasets. (Technical report)

SVM

▶ Support Vector Machine

SVM Supervector

An SVM (Support Vector Machine) is a two class classifier. It is constructed by sums of kernel function $K(.,.)$:

$$f(x) = \sum_{i=1}^{L} \alpha_i t_i K(x, x_i) + d \qquad (1)$$

t_i are the ideal outputs (-1 for one class and $+1$ for the other class) and $\sum_{i=1}^{L} \alpha_i t_i = 0 (\alpha_i > 0)$ The vectors x_i are the support vectors (belonging to the training vectors) and are obtained by using an optimization algorithm. A class decision is based upon the value of $f(x)$ with respect to a threshold. The kernel function is constrained to verify the Mercer condition: $K(x, y) = b(x)^t b(y)$, where $b(x)$ is a mapping from the input space (containing the vectors x) to a possibly infinite-dimensional SVM expansion space.

In the case of speaker verification, given universal background (GMM UBM):

$$g(x) = \sum_{i=1}^{M} \omega_i N(x, \mu_i, \Sigma_i), \qquad (2)$$

where, ω_i are the mixture weights, $N()$ is a Gaussian, and$(\mu_i + \Sigma_i)$are the means and covariances of Gaussian components. A speaker (s) model is a GMM obtained by adapting the UBM using MAP procedure (only means are adapted: (μ^s)). In this case the kernel function can be written as:

$$K(s_1, s_2) = \sum_{i=1}^{M} (\sqrt{\omega_i} \Sigma_i^{-(1/2)} \mu_i^{s_1})^t (\sqrt{\omega_i} \Sigma_i^{-(1/2)} \mu_i^{s_2}). \qquad (3)$$

The kernel of the above equation is linear in the GMM Supervector space and hence it satisfies the Mercer condition.

▶ Session Effects on Speaker Modeling

Sweep Sensor

It refers to a fingerprint sensor on which the finger has to sweep on the platen during the capture. Its capture area is very small and it is represented by few pixel lines.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Synthesis Attack

Synthesis attack is similar to *replay attack* in that it also involves the recording of voice samples from a legitimate client. However, these samples are used to build a model of the client's voice, which can in turn be used by a *text-to-speech synthesizer* to produce speech that is similar to the voice of the client. The text-to-speech synthesizer could then be controlled by an attacker, for example, by using the keyboard of a notebook computer, to produce any words or sentences that may be requested by the authentication system in the client's voice in order to achieve false authentication.

► Liveness Assurance in Face Authentication
► Liveness Assurance in Voice Authentication
► Security and Liveness, Overview

Synthetic Biometrics

► Biometric Sample Synthesis

Synthetic Fingerprint Generation

► Fingerprint Sample Synthesis
► SFinGe

Synthetic Fingerprints

► Fingerprint Sample Synthesis

Synthetic Iris Images

► Iris Sample Synthesis

Synthetic Voice Creation

► Voice Sample Synthesis

System-on-card

Smartcard has complete biometric verification system which includes data acquisition, processing, and matching.

► On-Card Matching

T

Tablet

▶ Digitizing Tablet

Tamper-Proof

The term tamper-proof refers to as a functionality of a device that enables the system to resist and/or protect itself from tampering acts. This functionality sometimes implemented as a combination of a self-destruction mechanism and sensors that detect any unauthorized access to the device including vandalism. This functionality is also known as temper-resistant or anti-tampering.

▶ Finger Vein Reader

Tamper-proof Operating System

Raul Sanchez-Reillo
University Carlos III of Madrid; Avda. Universidad, Leganes (Madrid), Spain

Synonyms

Malicious-code-free Operating System; Secure Biometric Token Operating System

Definition

Operating System with a robust design, as not to allow the execution of malicious code. Access to internal data and procedures are never allowed without the proper authorization. In its more strict implementations, this Operating System will have attack detection mechanisms. If the attack is of a certain level, the Operating System may even delete all its code and/or data.

Introduction

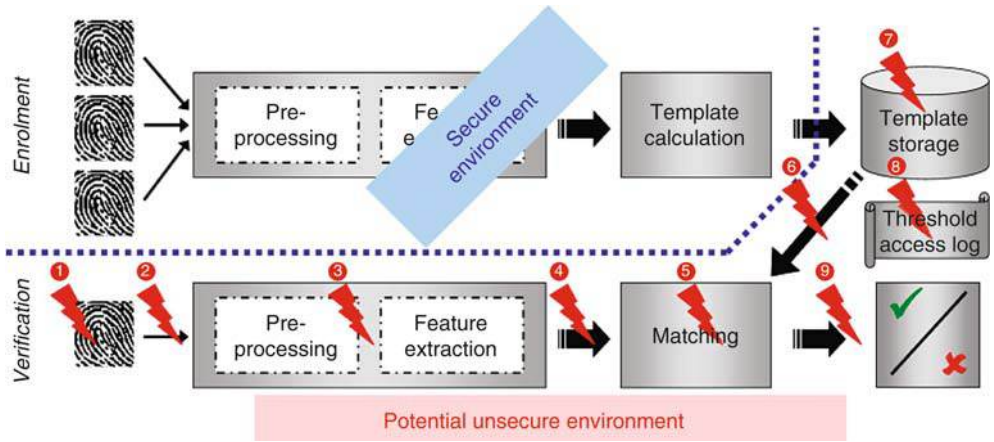
The handling of sensible data in Information Systems is currently very usual. Which data is to be considered sensible is up to the application, but at least we can consider those such as personal data, financial data as well as access control data. Actors dealing with such Information System (clients/citizens, service providers, integrators, etc.) have to be aware of the security level achieved within the system.

Although this is a very important issue in any system, when biometric information is handled it becomes a critical point. Reason for this is that biometric information is permanently valid, as it is expected to be kept the same during the whole life of a person. While a private key can be changed as desired and even cancelled, a user cannot change his fingerprint (unless changing finger) or even cancel it. If cancelling biometric raw data, the user will be limited, in case of fingerprints, to 10 successful attacks during his/her whole life. These kind of considerations has already been published even back to 1998, as it can be read in [1].

Therefore, biometric systems have to be kept as secure as possible. There are several Potential Vulnerable Points (PVPs) in any Biometric System, as it can be seen in Fig. 1. All those 9 PVPs have to be considered when designing a biometric solution. A good introduction to threats in a Biometric System can be found in [2, 3], and in BEM [4].

- PVP 1 has to deal with user attitudes, as well as capture device front end. Regarding user attitude, an authorised user can provide his own biometricsample to an impostor unknowingly, unwillingly, or even willingly. From the capture device

© 2009 Springer Science+Business Media, LLC



Tamper-proof Operating System. **Figure 1** Potential vulnerable points in a Biometric System where Enrolment is considered secured.

front-end point of view, such device may not be able to:

- Detect a nonlive sample
- Detect the quality of the input sample, being able to discard those under a determined threshold
- Protect the quality threshold against manipulation
- Detect degradation of its own degradation
- Resist environmental factors
- Eliminate residual information from previous captures
- Detect and discard sample injection
- Deny successive and fast sample presentation

- PVP 2 is directly related to the threat group 3 of BEM. It is basically focused on the capture device back-end, as well as the front-end of the Biometric Algorithm. Captured sample could be intercepted and/or reinjected, to provide a ▶ replay attack. Major problem relies on the potential lost of the user's biometric identity. Also, another threat is a ▶ hill-climbing attack by injecting successive biometric samples.
- PVPs 3, 4, 6, 7, and 8 could be treated as in any other IT system (Trojans, Viruses, communications interception, data injection, hill climbing attacks, etc.). So the same kind of study shall be done. It is in this kind of PVPs where a Tamper-proof Operating System can be of help. It is important to note that sensibility related to biometric related information, covers not only the sample data, feature vectors, and templates, but also thresholds, access logs, and algorithms.
- PVP 5, being also a typical point of study in any IT system, has here more importance depending on the information that could be given by the system after the matching. If matching result is not given just by an OK / ERROR message, but also carries information about the level of matching acquired, this could be used by an attacker to build an artificial sample, by hill-climbing techniques. For this PVP also, the Tamper-proof O.S. can play an important role.

Biometric Devices

Regarding Biometrics, a Tamper-proof Operating System is intended to be running in some (or any) elements which are part of the Biometric System. The idea of this kind of Operating Systems is not new, as they are already implemented in other areas, such as ▶ smart cards for financial services. This kind of electronic devices are designed under a basic security rule: "Not only the device has to work under its constrained conditions of user, but also has to stop working outside those conditions". In few words, this means that, for example, if the smart card is expected to work with a supply voltage from 4.5 to 5.5 volts, it does not have to work outside the range (e.g., if supply voltage is 4.4 or 5.6, not even a response has to be obtained from the card). Related to the Operating System inside the card, this covers things like not allowing the execution of any undefined/undocumented command, or not

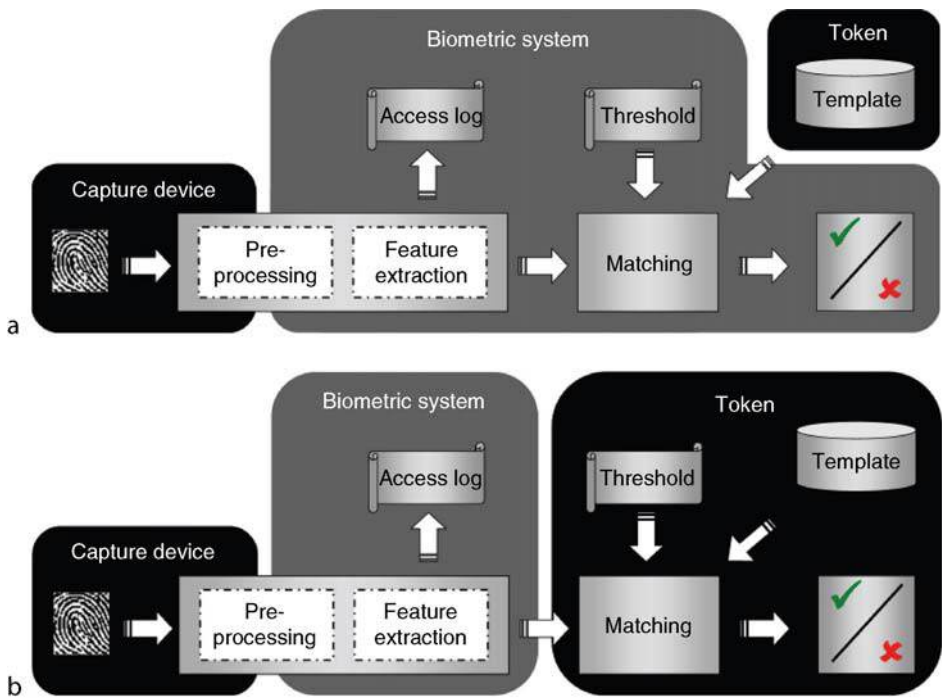
being able to install new functions that can behave as Trojan Horses or Viruses.

With this example, the reader can think that this kind of products does not really exist, because several papers have been published related to security problems with smartcards (e.g., [5] and some general audience press). It has to be stated that not all times an integrated circuit identification card is referred, it is really a smartcard (i.e., a microprocessor-based identification card with a tamper-proof O.S.). Also, some real smartcards have not been properly issued, leaving some critical data files unprotected, or not using the security mechanisms provided. Rules to be followed to properly use a smartcard can be found in [6].

This same kind of rules can be applied to all kind of biometric devices. Obviously, depending on the system architecture, biometric devices can be of very different kinds. Figure 2 shows two possible architectures of a biometric authentication system, which are usually known as (a) match-off-card (also known as match-off-token), and (b) match-on-card (or match-on-token). Apart from these two, many other schemes can be designed. The term "match" should be changed for "Comparison". So instead of "match-off-card", "off-card biometric comparison" should be used. But within text "match-off-card" and "match-on-card" are used due to being terms widely used among the industry.

In a match-off-card system (e.g., [7]), we can consider a simplification of the system as composed of three devices: the capture device, the token or card where the user's template is stored, and the rest of the system, which will be named as "Biometric System". Major difference with the match-on-card system (e.g., [8]) is that here the token only stores the user's template, while in the match-on-card version it also performs some biometric-related computations.

In any case, any of those devices should be designed following the rules given below regarding a Tamper-proof O.S. Being this viable, if those devices are developed as embedded systems, major problems can arise when one of those devices (typically the Biometric System) is running on a general purpose computer, where little or no control is available for installed applications and data exchange.



Tamper-proof Operating System. **Figure 2** Some architectures of biometric authentication systems, splitting tasks in several devices.

Requirements for a Tamper-Proof O.S.

Once focused the environment where a Tamper-proof O.S. has to be found in Biometric Devices, it is time to start its design. A good starting point will be following all previous works dealing with smart cards. The reason for that is to transfer the know–how of near 30 years of secure identification tokens given by the smart card industry [9]. This same ideas can be extrapolated to other biometric devices, not only personal tokens.

First thing to consider when designing a Tamper-proof O.S. is the different life phases that the biometric device will have. All devices, specially those related to personal authentication, should go through different life stages, from manufacturing to its use by the end users. As information handled by them is really sensible, extra protection should be taken to avoid robbery, emulation, or fraudulent access to the device or its information. Therefore, security mechanisms will be forced in each life stage. Those mechanisms are mainly based on Transport Keys, which protect access to using the device in each change of its life phase. Life phases defined are:

- Manufacturing: where the device is assembled. The microcontroller within the device should be protected by a Transport Key, before delivered to the next stage. The way to compute that Transport Key for each microcontroller, will be sent to the company responsible of the next phase by a separate and secured way.
- Personalization: In this phase, each device is differentiated from all others by storing some unique data related to the final application, user, and access conditions. Some times this phase is split in several subphases, specially when the device has to be personalized for the application (prepersonalization) and then for the final user (personalization), as it may happen with identification tokens. In this phase, also Data Structure regarding the applications may be created, as well as the full security architecture.
- Usage: The end user is ready to use the device.
- Discontinuation: Due to ageing, limited time use, accidents, or attack detection, the device may be out of use. This can be temporary (for example, when keys are blocked), or permanent (no re-activation is allowed). It has to be guaranteed that once discontinued, such device shall not be able to be used.

Entering in details regarding the requirements for a tamper-proof operating system, we can state the following general rules:

- Mutual Authentication mechanisms have to be used before exchanging any kind of biometric information. In any communication, both parts have to be sure that the other party is a reliable one.
- To avoid ► replay attacks, some time-stamping-like mechanisms have to be used (e.g., generation of session keys to sign/cipher each message exchanged).
- Only the manufactured designed commands can be executed. No possibility of downloading new commands has to be allowed. Therefore, flash reprogramming and device updating are strongly discouraged.
- Before executing anything in the biometric device, full integrity check (both cryptographic and semantic) of the command and its data has to be performed. Some attacks would try to exploit undefined cases in the semantics of a command exchanged.
- All sensible data (sample data, feature vectors, templates, and thresholds) has to be transmitted ciphered.
- If there is a command related to changing parameters, it has to be sent with all security mechanisms allowed, as the system can be even more vulnerable to attacks related to changing those parameters (e.g., quality or verification thresholds).
- Feedback information from the device to the external world has to be as short as possible to avoid hill-climbing attacks. For example, a device performing comparisons in an authentication system has to provide only a YES/NO answer, but not giving information on the matching score obtained.
- Attack detection mechanisms have to be considered. If an attack is detected, then the device has to stop working, and a reinitialization has to be made. If the detected attack is consider extremely serious, the device may consider deleting not only all temporal data, but also its permanent data or even it programming code.
- Successive failed attempts to satisfy any security condition has to be considered as an attack, and therefore, the device has to be blocked, as it happens with a PIN code in a smartcard.
- No direct access to hardware resources (e.g., memory addresses, communication ports, etc) can be

allowed. Most virus and Trojan horses benefit for not following this rule.
- As soon as data is no longer needed by the Operating System, it has to be erased as to prevent latent data to be acquired in a successful attack.

Most of these requirements can be satisfied by defining a security architecture based on cryptographic algorithms. Several implementations can be followed. If the developer is not familiar with these mechanisms, it is suggested to follow the secret codes/secret keys architecture of a smartcard, and the Secure Messaging mechanism [6, 9]. These can be directly applied to personal Tokens, and upgraded to other kind of biometric devices.

Example of an O.S. Instruction Set

When implementing a Tamper-proof O.S. several design decisions have to be made: Frame formats, time-outs, number of retries, etc. All these issues depend on the communication strategy followed by the whole biometric system. Therefore, no general rule can be given to the designed.

Regarding the instruction set, a minimal list of functions can be considered, depending on the device where the O.S. is to be included. This is also dependent on the platform chosen. As an example, the instruction set for a limited-resources platforms is given. This instruction set has been proposed to ISO/IEC JTC1/SC37 to be considered as a lighter version of BioAPI, the standardized Application Program Interface for biometric applications. This lighter version is called BioAPI Lite and is being standardized as ISO/IEC 29164.

Commands needed by a limited biometric device, depends on the functionality of such device. Obviously is not the same a capture device, than a personal token. But in general terms these commands can be classified in four major groups: Module Management, Template Management, Biometric Enrolment, and Biometric Process.

Management Commands relate to manage the overall module behavior. Four commands can be considered in this group:

- Initialize: Tells the module to initialize itself, opening the offered services and initialize all security for ciphered data exchange. This command is to be called any time a session is started (power on, session change, etc). Without being called, the rest of the commands shall not work.
- Close: Tells the module to shutdown.
- Get Properties: Provides information on capabilities, configuration, and state.
- Update Parameters: Updates parameters in module. One of such parameters can be the comparison threshold. For that reason, this function is recommended to be used with all security mechanisms available.

Template Management Commands refer to those functions needed to store and retrieve templates from the module. These functions will be supported by those modules that are able to store users' templates. These set of commands are expected to be used by personal tokens or small databases. The functions defined are:

- Store Template: Stores the input template in the internal biometric module database.
- Retrieve Template: Obtain the referenced template from the biometric module.

The next group is the Biometric Enrolment Commands. This group of functions will be considered in systems where enrolment is to be made internally. Due to the different process of enrolment, even for a single biometric modality (e.g., different number of samples needed), in limited devices a multi step procedure is suggested. First, user will call functions related to obtain samples for the enrolment and then a call to the Enrol function will have to be done. Commands defined are:

- Capture for Enrol: Performs a biometric capture (using onboard sensor), keeping the information in module for later enrolment process. The number this function is called depends on the number of samples the module needs to perform enrolment. As this operation involves user interaction, biometric module manufacturer shall consider time-out values to cancel operation, reporting that situation in the Status code returned.
- Acquire for Enrol: Receives a biometric sample to keep the information in module for later enrolment process. The number this function is called depends on the number of samples the module needs to perform enrolment. Depending on module capabilities, input data can be a raw sample, a preprocessed one, or its corresponding feature vector.

- Enrol: Performs an enrolment to create a template and stores the template in module. To execute this function, either Capture for Enrol or Acquire for Enrol functions has to be called in advance. Enrol with process with the samples temporally stored in the module. The return value is the number of template internally assigned.
- Erase Enrolments: Erases all enrolment templates or the indicated (by number) template.

Finally, the fourth group is dedicated to all those commands that are dealing with biometric functions. It covers the capture process, feature extraction, and comparison. Even with comparison, it handles comparisons with internal templates, or templates coming from the external world.

- Capture: Performs a biometric capture (using on-board sensor), returning biometric sample.
- Process: Processes biometric sample to create comparable recognition data (feature vector). Depending on module capabilities, the input sample can be a raw sample or a preprocessed one.
- Capture and Process: Performs a biometric capture (using on-board sensor), returning its feature vector.
- Compare External: Compares a feature vector with the template sent by the external world.
- Process and Compare External: Processes a biometric sample and compares it with the template sent by the external world.
- Capture and Compare External: Perform a biometric capture (using on-board sensor), process the biometric sample, and compares it with the template sent by the external world.
- Compare Internal: Compares a feature vector with templates stored in the module. If the input parameter is 0xFF, comparison will be done with all templates stored. In other case, comparison is done only with the template whose internal number is given at the input parameter.
- Process and Compare Internal: Processes a biometric sample and compares it with templates stored in the module. If the input parameter is 0xFF, comparison will be done with all templates stored. In other case, comparison is done only with the template whose internal number is given at the input parameter.
- Capture and Compare Internal: Perform a biometric capture (using on-board sensor), process the biometric sample, and compares it with templates stored in the module. If the input parameter is 0xFF, comparison will be done with all templates stored. In other case, comparison is done only with the template whose internal number is given at the input parameter.

Some of these instructions involve user interaction. Therefore, manufacturer shall consider time-out values to cancel operation if it is exceeded, reporting that situation within the protocol used.

Applicability of Tamper-Proof O.S.

As mentioned above, this kind of Operating System is desirable to be included in all devices related to Biometric Identification, but unfortunately this is not always possible. As in many applications a general purpose computer is used, general purpose Operating Systems are used (such as Windows, Linux, etc.). Developing those O.S. in a Tamper-proof way, without restricting usability and generality is nearly impossible. Therefore, Tamper-proof Operating Systems are meant for those embedded systems, sensors and personal tokens, dealing with personal identification.

Using this kind of Tamper-proof O.S. in these devices, restrict the number of security holes to the minimum within the device, and to be concentrated only in those general purpose systems used. As some tasks will be performed in such secured devices, security leaks will be avoided. For example, if a biometric system uses personal tamper-proof tokens with match-on-card capability, the user's template will never be exposed, and possibility of hill-climbing or replay attacks will be cancelled. Thus, all comparison and decision blocks will be secured, restricting the potential security problems to the relevant previous modules.

Summary

Due to the sensibility of biometric data, security in biometric devices has to be considered. One of the ways to protect privacy is to include a Tamper-proof operating system. This O.S. would not allow direct access to hardware resources of the device, neither to temporary nor permanent data. This O.S. has also to control the different life stages of the device. A set of requirements have been defined that have to be considered when developing such Tamper-proof O.S.

Finally an example of the commands to be covered by some devices have been given. Including this kind of O.S. in all biometric devices will improve the security of the whole system. Unfortunately, when some parts of the biometric system has to be implemented in a general purpose computer with an open Operating System, applying these rules is not easy.

Related Entries

- ▶ Biometric Security Threat
- ▶ Biometric Token
- ▶ Biometric Vulnerabilities
- ▶ Match-off-Card
- ▶ Match-on-Card
- ▶ Template Security

References

1. Rejman-Greene: Security considerations in the use of biometric devices. Information Security Technical Report **3**, 77–80 (1998)
2. Ratha, N.K., Connell, J.H.B.R.M.: Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal **40**(3), 614–634 (2001)
3. Roberts, C.: Biometric attack vectors and defences. Computers & Security **26**(1), 14–25 (2007)
4. Criteria, C.: Biometric evaluation methodology supplement (bem). Common Methodology for Information Technology Security Evaluation - *http://www.cesg.gov.uk/site/ast/biometrics/media/BEM_10.pdf* (2002)
5. Matthews, A.: Side-channel attacks on smartcards. Network Security **2006**(12), 18–20 (2006)
6. Sanchez-Reillo, R.: Achieving Security in Integrated Circuit Card Applications: Reality or Desire? IEEE Aerospace and Electronic Systems Magazine **17**, 4–8 (2002)
7. Sanchez-Reillo, R., Gonzalez-Marcos, A.: Access control system with hand geometry verification and smart cards. Aerospace and Electronic Systems Magazine, IEEE **15**(2), 45–48. (2000). DOI 10.1109/62.825671
8. Sanchez-Reillo, R.: Smart card information and operations using biometrics. Aerospace and Electronic Systems Magazine, IEEE **16**(4), 3–6 (2001). DOI 10.1109/62.918014
9. *ISO/IEC_JTC1/SC17*: ISO/IEC 7816 Parts 3, 4, 8, 9 & 11 (1987–2005)

Target Detection

- ▶ Human Detection and Tracking

Target Population

- ▶ Test Sample and Size

Target-Dependent Fusion

- ▶ Fusion, User-Specific

Technology Tests

Technology tests are those in which biometric algorithms enroll and compare archived (i.e., previously-collected) data. An essential characteristic of technology testing is that the test subject is not "in the loop" – the test subject provides data in advance, and biometric algorithms are implemented to process large quantities of test data. Technology tests often involve cross-comparison of hundreds of thousands of biometric samples over the course of days or weeks. Methods of executing and handling the outputs of such cross-comparisons are a major component of technology-based performance testing standards. Technology tests are suitable for evaluation of both verification- and identification-based systems, although most technology tests are verification-based.

- ▶ Performance Testing Methodology Standardization

Template

The features extracted from an individual's biometric trait during enrollment is stored in the biometric database and is referred to as a template. During authentication, the system compares an individual's biometric features against this template. The template

may be updated over time in order to reflect changes in an individual's trait (if any).

► Biometrics, Overview
► On-Card Matching

Template Distortion

Template distortion consists of applying either invertible or non invertible distortions to a biometric template. The distortion can be performed using either an invertible or a non invertible transform. In both cases the transform is chosen on the basis of a user key, which must be known when authentication is performed. In the case when an invertible transform is chosen, the whole security of the system relies on the key. On the contrary when non invertible transforms are used, even if the key is known by an adversary, no significant information about the template can be acquired. Then, the distorted data are stored in the database. Different distorted data can be generated from the same original data, simply by changing the transform key. Moreover, even if the database is compromised, the biometric data cannot be retrieved unless user dependent keys are revealed, when dealing with invertible transforms.

► Conformance Testing for Biometric Data Interchange Formats, Standardization of
► Iris Template Protection

Template Protection

It is a method to keep away from biometric template attack, storing transformed template data rather than raw template data in a biometric system.

► Iris Template Protection
► User Interface, System Design

Template Reconstruction

► Template Security

Template Security

Andy Adler[1], Raffaele Cappelli[2]
[1]Carleton University, Ottawa, ON, Canada
[2]Biometric System Laboratory - DEIS - University of Bologna, Italy

Synonyms

Image regeneration from templates; Template reconstruction

Definition

Template security refers to techniques which allow regeneration of enrolled images from templates. Such image regeneration poses a security and privacy vulnerability because the images may be used to spoof or masquerade as the enrolled individual. Image regeneration is of two types. The first relies on decoding the features in the template and estimating a biometrically reasonable image with the appropriate features. Results have been published for fingerprint templates, but such algorithms are easily implemented for face and iris recognition. The second type of image regeneration uses the ability to compare images against the target and obtain the match score to perform hill-climbing to iteratively improve an image estimate. Appropriate biometric template security measures requires strong encryption of all biometric data, including templates and match results.

Introduction

It is generally understood that source biometric images – those captured at the sensor for enrollment and matching – need to be handled carefully; compromise of these images can impact user privacy and provide a pathway for security breaches (see ► Biometric Vulnerabilities, Overview). On the other hand, other types of biometric data are not necessarily managed with the same care as the biometrics images, and recent work has shown that it is often possible to regenerate the source images, or learn other important information, from these data. Image regeneration may be performed from two data sources, biometric

templates and match (or similarity) score results. This topic has come to be known as "biometric template security," or "template reconstruction," even when it applies to security issues around the match scores. Perhaps the best overview of the issues in biometric template security is given by Jain et al. [1, 2].

Biometric templates carry the most important biometric information – the features considered most discriminating for the identity of the subject. and thus present an important concern for privacy and security of systems. The basic concern is that templates may be used to spoof the owner of the document. Biometric algorithm vendors have largely claimed that it is impossible or infeasible to regenerate the image from the templates [3]; thus biometric templates are sometimes considered to be effectively non-identifiable data, much like a password hash. These claims are supported by: (1) the template records features (such as fingerprint minutiae) and not image primitives, (2) templates are typically calculated using only a small portion of the image, (3) templates are small – a few hundred bytes – much smaller than the sample image, and (4) the proprietary nature of the storage format makes templates infeasible to "hack." Two pathways are considered from which to regenerate images from templates: (1) from the template directly, based on a knowledge of the features, and (2) from match score values from a biometric algorithm.

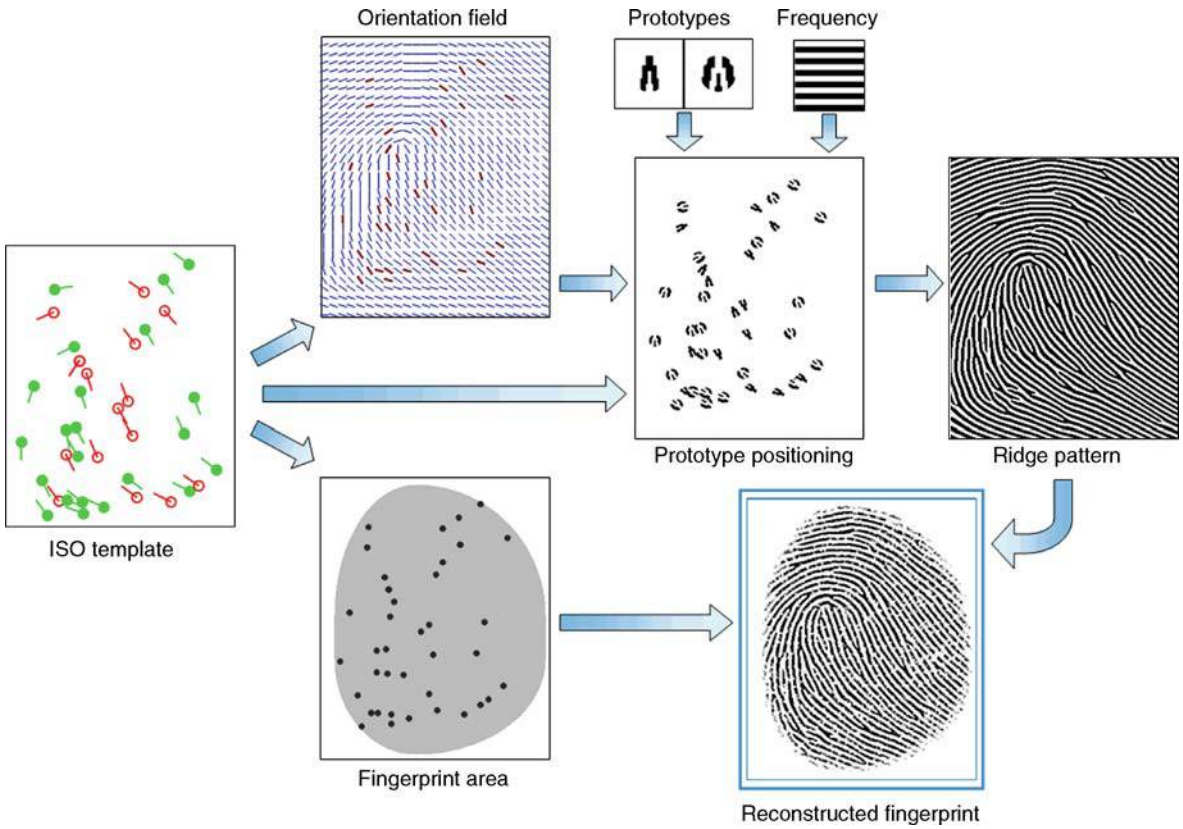
Image Regeneration from Templates

The goal of image regeneration from a biometric template is to compute an image which best matches the feature values encoded in the template, while maintaing a "reasonable" appearance. The constraints on the image appearance depends on whether the regenerated images are designed to be shown to humans (experts or casual observers) or only to be used to fool a biometric algorithm. In order to regenerate images in this way, it is necessary for templates to be available in unencrypted form. Thus, encryption of template data storage does impede this vulnerability; however, since templates must be available in unencrypted form to perform matching, they are vulnerable at that point.

Published work on image regeneration from templates is for fingerprints, for the reason that regeneration is trivial for most iris and face recognition templates, in which the template features are typically based on subspace image transforms. A biometric template based on such transforms may be used to regenerate the image as follows: assume the template feature vector, $\mathbf{y}$, is computed from an image, $\mathbf{x}$ using a transform that can be approximated by a linear equation $\mathbf{y} = \mathbf{H}\mathbf{x}$ for a linear subspace transform expressed as a matrix, $\mathbf{H}$, then a reconstructed image, $\hat{\mathbf{x}}$, can be computed from $\hat{\mathbf{x}} = \mathbf{H}^{\dagger}\mathbf{y}$ using a pseudoinverse $\mathbf{H}^{\dagger}$. Construction of $\mathbf{H}^{\dagger}$ would use the well understood techniques of inverse problem theory (e.g., [4]). Images reconstructed in this way would typically suffer some blurriness, due to the inherent ill-conditioning of such inverse problems.

Because fingerprint template features cannot be expressed as linear functions of the image data, reconstruction of fingerprint images from templates is a nonlinear, and considerably more difficult problem. The earliest template reconstruction technique for fingerprints was proposed by Hill [5], who developed an ad-hoc approach to calculate an image from the template of an unspecified fingerprint system vendor. Software was designed to create line pattern images which had a sufficient resemblance to the underlying ridge pattern to be verified by the match software. This work also devised a simple scheme to predict the shape (class) of the fingerprint using the minutiae template. The algorithm iterated over each orientation, core and delta position keeping the image with the best match score. It is worth noting the reconstructed line patterns do not visually resemble a fingerprint, although these images could be easily improved manually or automatically.

More recently, Ross et al. [6, 7] have demonstrated a technique to reconstruct fingerprint images from a minutiae description, without using match score values. First, the orientation map and the class are inferred based on analysis of local minutiae triplets and a nearest neighbor classifier, trained with feature exemplars. Then, Gabor-like filters were used to reconstruct fingerprints using the orientation information. Correct classification of fingerprint class was obtained in 82% of cases, and regenerated images resembled the overall structure of the original, although the images were visually clearly synthetic and had gaps in regions which lacked minutiae. Another valuable contribution of this work is calculation of the probability density fields of minutiae; such information could be used to attack fingerprint based biometric encryption schemes.



Template Security. **Figure 1** A functional schema of the reconstruction from fingerprint templates using the approach of Cappelli et al. [9]. The minutiae information in the template is used to: estimate the orientation field, generate a reasonable fingerprint area and place synthetic minutia prototypes on an blank image. Then, given a constant frequency, a ridge-line pattern is generated according to the orientation field, starting from the prototypes. Finally noise is added to make the reconstructed image more realistic.

In 2007, Cappelli et al. [8, 9] demonstrated a technique which allows synthesis of highly realistic fingerprint images from minutiae data in templates. The reconstruction approach is based on a sequence of steps to estimate various aspects of the original unknown fingerprint from the template (Fig. 1): the fingerprint area, the orientation field, and the ridge-line pattern. First, a simple elliptical model is adopted and its parameters estimated by calculating the minimal area that encloses all the minutiae. Next, starting from each minutia direction, the orientation field is estimated by optimizing the parameters of the model proposed in [10]. Finally, the ridge-line pattern is generated, starting from the estimated orientation field estimated and from generic minutiae prototypes positioned according to the template information. The reconstructed images are very similar to the original fingerprints. Although the reconstructed images may not fool a human expert, they may be used to successfully attack automatic recognition systems; the percentage of successful attacks against nine different systems was 81% at a high security level, and 90% at a medium security level.

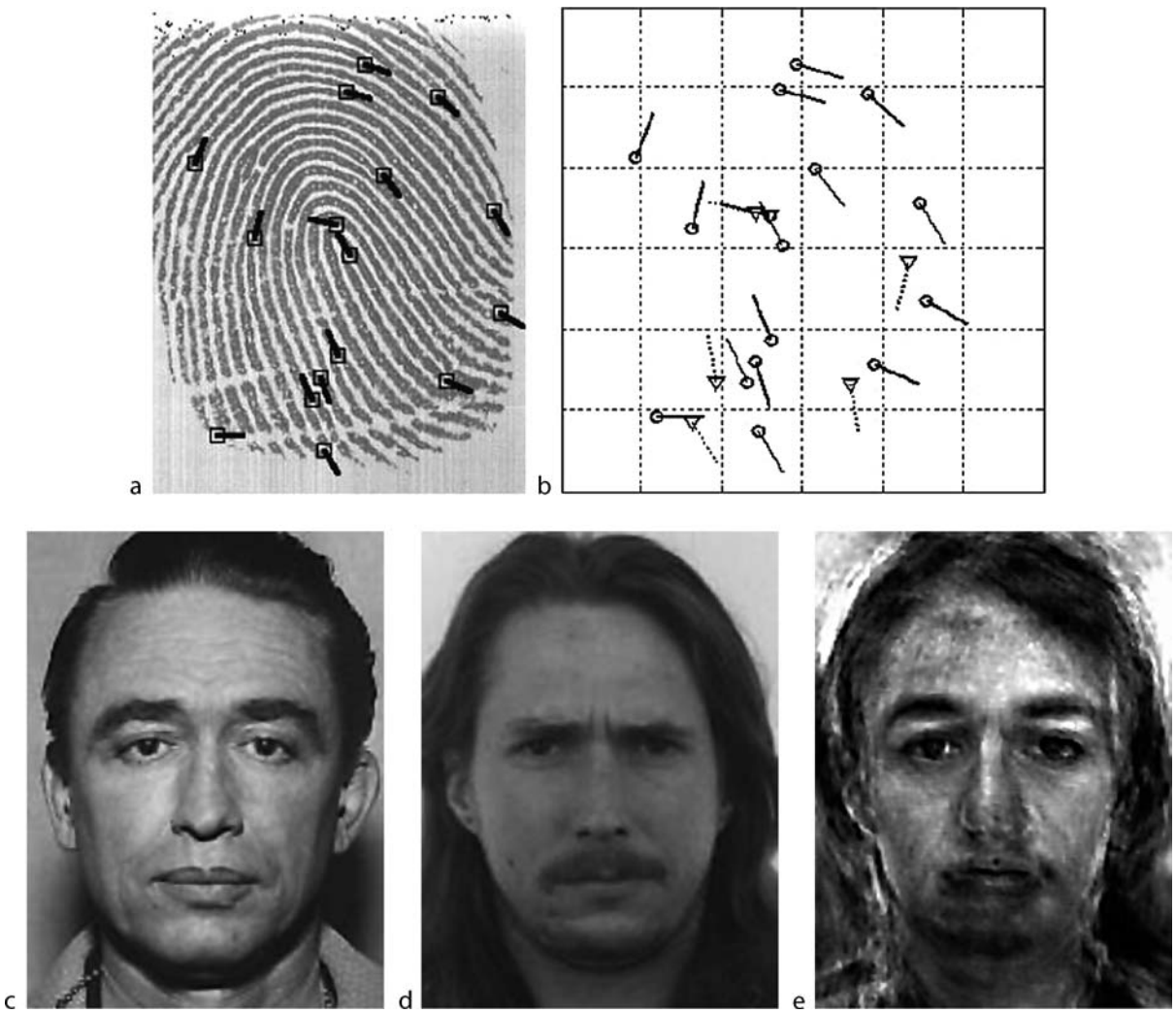
Image Regeneration from Match Scores

The other approach to regeneration of biometric images is based on match score values and does not require access to the template. This means that template encryption is not a countermeasure, but also means that a large number of match score comparisons are needed. For this to be feasable, the requirements are

the ability to present arbitrary images for matching against a target, and access to calculated match scores. The goal is to: (1) determine an image which matches against the target for the specific biometric algorithm, and (2) determine a good estimate of the original image. Clearly, if one can test arbitrary images, one could mount a brute force attack. Given a biometric database of sufficient quality and variety, it should be possible to attain the first goal in approximately 1/FAR attempts. A brute-force attack would be guaranteed to succeed in the second goal, but the size of image space is extremely large.

Brute force searches would only be necessary if biometric image space were random, and nothing could be learned from the output of previous tests. Soutar et al. [11] first proposed the possibility of "hill-climbing" in order to practically regenerate images from match score data (Fig. 2). A hill-climbing algorithm functions as follows:

1. *Initial image selecion*: Choose an initial image estimate (*IM*). Typically, a sample of initial biometric patterns are tested and the one with the largest match score, *MS*, is selected.



Template Security. Figure 2 Regenerated images using hill-climbing techniques. (**a**–**b**): Regenerated fingerprint minutiae (from [13]). The target fingerprint with labeled minutiae (**a**), and regenerated minutiae positions (**b**). (**c**–**e**): Regenerated face images (from [12]). The target face image (**c**); the initial selected image for hill-climbing (**d**), and regenerated face image (**e**).

2. *Iterative estimate improvement*:
   a. Modify *IM* (to get IM_{test}) in a random, but biometrically reasonable way (details below).
   b. Calculate MS_{test} for IM_{test}.
   c. If $MS_{test} > MS$, set $IM = IM_{test}$ and $MS = MS_{test}$.
   d. End iterations if *MS* is no longer increasing.

The only difficulty to a practical implementation of this algorithm is to implement "biometrically reasonable" modifications. For face images, Adler [12] added a small factor times a PCA (eigenface) component to the face image. For fingerprint minutiae, Uludag and Jain [13] made modifications to perturb, add, replace, or delete an existing minutiae point at each step. The key constraint is that such modifications attempt to maintain "biometric feasibility" in the search space. Other image modifications, such as changing random pixels in the image, do not converge under hill-climbing.

In fact, "hill-climbing" algorithms are simply one type of multidimensional optimization algorithm. Other methods for unconstrained minimization (or maximization) such as the Nelder–Mead simplex appear to perform equally or better than hill-climbing (unpublished observations).

In order to protect against regeneration of biometric images, Soutar et al. [11] suggested that match score output be quantized to a limited set of levels. The idea is that small image modifications are unlikely to push the *MS* up by one quantum, so that the hill-climbing algorithm will not see the effect of its changes. This recommendation is maintained in the BioAPI specification [14]. However, by an appropriate modification of the algorithm, Adler showed that hill-climbing could still function [15]. Each hill-climbing iteration is applied to a quadrant of *IM*. Before each calculation, noise is added to the image in the opposite quadrant, in order to force the match score to a value just below the quantization threshold. This means that the quantized match score is brought into a range where it provides useful information. Images were successfully regenerated for quantization levels equal to a 10% change in FAR.

These results suggest that biometric images can generally be regenerated if: (1) arbitrary images can be input into the biometric system, and (2) raw or quantized match score values are output. The images calculated are of sufficient quality to masquerade to the algorithm as the target, and give a good visual impression of the biometric characteristics. In order to prevent this attack, it is necessary to either limit image input, or to provide only Match/Nonmatch decisions.

Consequences of Template Security Breaches

There has been some criticism of the research on biometric template security which accuses the authors of fearmongering (need ref). Specifically, it is claimed that there are no (or very few) practical scenarios in which image regeneration from templates is a serious security consideration. While this criticism has some merit, there do exist several security breaches which are facilitated by these techniques, such as the following examples:

- *Fraudulent use of passports*. Passports conforming to the most recent ICAO specification (and required for visa–waiver entry into USA) encode the fingerprint or iris templates of the holder in an embedded smart card. Similarly, the new ILO standard for the seafarers ID card [16] encodes the fingerprint minutiae into a 2D barcode on the document in a standardized format. In general, it may be assumed that even when these data are encrypted, it will be possible to decode, either through errors in the issuance process (e.g., [17]) or by using the public keys of the issuance agencies. After decoding the document data, the template may be read, and an image reconstructed which is sufficiently similar to match. It would then be possible to fabricate a spoofed biometric to allow a criminal or terrorist to fraudulently use the document and bypass the biometric security.
- *Biometrically locked digital media*. Biometrics have been proposed as a way to control access to digital media, with the primary interest being in preventing copyright infringement. Digital documents encoded with the biometric of the user(s) with approved access will presumably be subject to attacks, especially since both the documents and the software to access them will be widely distributed. The techniques described in this article would facilitate attacking the digital locks and impact the privacy of the content owner.
- *Regenerating watchlist images from match scores*. In many cases, governments may allow collaborating agencies to perform searches against a biometric watch list; however, for security reasons, the

primary agency may not want to distribute watch list images. Using the techniques described here, it may effectively be possible to generate these watchlist images from the match score data, thus bypassing the security.

Summary

This article summarizes the current research in template security. This work shows that, in all cases tested, a high quality image of an enrolled fingerprint or face can be regenerated if access is given to biometric templates or to match scores. This is strong evidence to refute the somewhat naïve assumption commonly made that biometric templates are secure in a similar way to a cryptographic hash function. Based on these results, a prudent design for biometric security should consider *any* biometric data to potentially "leak" information about the source images, and provide a potential attack pathway. One partial solution is the use of cryptographic techniques to protect biometric data in databases and communicated over networks.

Related Entries

- ▶ Biometric Encryption
- ▶ Biometric System Design
- ▶ Biometric Vulnerabilities
- ▶ Security and Liveness, Overview

References

1. Jain, A.K., Nagar, A., Nandakumar, K.: Biometric template security. EURASIP. J. Adv. Signal Proc. (2008). doi:http://hindawi.com/RecentlyAcceptedArticlePDF.aspx?journal=ASP&number=579416
2. Jain, A.K., Ross, A., Uludag, U.: Biometric template security: Challenges and solutions. In: Proceedings of 13th European Signal Processing Conference (EUSIPCO2005), Antalya, Turkey, 2005
3. International Biometric Group.: Generating images from templates. (2002) www.biometricgroup.com/reports/public/reports/templates images.html
4. Tarantola, A.: Inverse problem theory and methods for model parameter estimation. Society for Industrial and Applied Mathematics (2005). ISBN 0-89871-572-5
5. Hill, C.: Risk of Masquerade arising from the storage of biometrics. Dissertation, Austrailian National University (2001)
6. Ross, A., Shah, J., Jain, A.K.: Towards reconstructing fingerprints from minutiae points. In: Conference of SPIE Biometric Technology for Human Identification, II, **5779,** 68–80 (2005)
7. Ross, A., Shah, J., Jain, A.K.: From template to image: reconstructing fingerprints from minutiae points. IEEE Trans. Pattern Anal. Machine Intel. **29**, 544–560 (2007)
8. Cappelli, R., Lumini, A., Maio, D., Maltoni, D.: Evaluating minutiae template vulnerability to masquerade attack. In: Proceedings of 5th IEEE Workshop Auto Ident. Adv Technol. Alghero, Italy, 7–8 June 2007
9. Cappelli, R., Lumini, A., Maio, D., Maltoni, D.: Fingerprint image reconstruction from standard templates. IEEE Trans. Pattern Match. Mach. Intell. **29**, 1489–1503 (2007)
10. Vizcaya, P., Gerhardt, L.: A nonlinear orientation model for global description of fingerprints. Pattern Recogn. **29**, 1221–1231 (1996)
11. Soutar, C., Gilroy, R., Stoianov, A.: Biometric system performance and security. In: Proceedings of the Conference of IEEE Auto Identification Advanced Technology (1999)
12. Adler, A.: Sample images can be independently restored from face recognition templates. In: Proceedings of Can. Conf. Elec. Comp. Eng. Montréal, Canada, pp. 1163–1166 (2003)
13. Uludag, U., Jain, A.K.: Attacks on biometric systems: a case study in fingerprints. In: Proc. SPIE-EI 2004, Security, Steganography and Watermarking of Multimedia Contents VI **5306,** 622–633 (2004)
14. BioAPI Consortium.: BioAPI Specification version 1.1 (2001)
15. Adler, A.: Images can be regenerated from quantized biometric match score data. In: Proceedings of Can. Conf. Elec. Comp. Eng. Niagara Falls, Canada, pp. 469–472 (2004)
16. International Labour Organization.: Biometric testing campaign report (Addendum to Part I). Geneva (2005)
17. The Guardian.: Cracked it! (17 Nov. 2006)

Temporal Characterization of Faces

Temporal characterization of faces is the process of modeling the way appearance of a face varies with time. As a face moves, its appearance as captured in a video changes due to change in pose, illumination conditions, expression, etc. The pattern of these variations is often specific to the individual, containing distinguishing information that can be used by video-based face recognition systems for improved performance.

▶ Face Recognition, Video-based

Temporal Domain

When a face is captured over a period of time, as in video recording, it is often said that a facial image is available in temporal domain or that it has temporal resolution. In contrast, when only a single image of a face is available, as in a passport photograph, it is said that facial image is not available in temporal domain. In sensing data, a natural tradeoff is observed: either sensory data are of high spatial resolution or temporal resolution, but not of both at the same time. For example, an image of a face in a printable document is of high resolution, whereas faces observed live on TV are normally of very small resolution. As demonstrated by biological vision systems, recognizing an object that is observed in temporal domain (e.g., recognizing a face on TV) can be done just as efficiently or even more efficiently than recognizing the same object from a single high-resolution sample. For automated recognition systems however this is not the case yet.

► Face Databases and Evaluation

Tenprint Capture

► Biometric Technical Interface, Standardization

Tensor

In mathematics, tensor is a topic in multilinear algebra. A tensor, which can be expressed as a multidimensional array, is an object extending the notion of scalar, vector, and matrix. For example, tensor singular value decomposition (SVD) is a generalization of matrix SVD to multilinear higher-order SVD.

► Face Sample Quality

Test Sample and Size

Michael E. Schuckers
St. Lawrence University, Canton, NY, USA

Synonyms

Crew designs; Sample size; Target population

Definition

The testing and evaluation of biometrics is a complex task. The difficulties in such an endeavor include the selection of the number and type of individuals that will participate in this process of testing. Determining the amount of data to be collected is another important factor in this process. Choosing an appropriate set of individuals from which to collect biometrics data is another important aspect of testing a biometrics system.

Introduction

The assessment of a biometric system's matching performance is an important part of evaluating such a system. A biometric implementation is an ongoing process and as such will be treated as a process in the sense of Hahn and Meeker [1]. Thus, any inference regarding that process will be analytic in nature rather than enumerative as delineated by Deming [2]:

An enumerative study has for its aim an estimate of the number of units of a frame that belong to a specified class. An ► analytic study has for its aim a basis for action on the cause-system or the process, to improve product of the future.

Here focus is on determining the amount and type of data necessary for assessing the current matching performance of a biometrics system.

The matching performance measures that are commonly considered most important are the false match rate (FMR) and the false non-match rate (FNMR). One of the important parts of designing a test of a biometrics system is to determine, prior to completion, the amount of testing that will be done. Below calculations that explicitly allow for determining the

amount of biometric data which will be sampled are described. As with any calculations of this kind it is necessary to make some estimates about the nature of process beforehand. Without these, it is not possible to determine the amount of data to collect. These sample size calculations will be derived to achieve a certain level of sampling variability. It is important to recognize that there are other potential sources of variability in any data collection process.

Selection of the individuals from whom these images will be taken is another difficulty because of the need to ensure that the biometric samples taken are representative of the matching and decision making process. The goal of any data collection should be to take a sample that is as representative as possible of the process about which inference will be made. Ideally, some probabilistic mechanism should be utilized to select individuals from a targeted population. In reality, because of limitations of time and cost, this is a difficult undertaking and often results in a ► convenience sample, Hahn and Meeker [1].

Test Size Calculation

Determining the amount of biometric information to collect is an ongoing concern for the evaluation of a biometrics system. Several early attempts to address this problem include those by Wayman [3] and [4] as well as the description in Mansfield and Wayman [5] of the "Rule of 3" and the "Rule of 30". The former is due to several authors including Louis [6] as well as Jovanovic and Levy [7], while the latter, the so-called Doddington's Rule, is due to Doddington et al. [8]. Mansfield and Wayman note that *neither* of these approaches is satisfactory since they assume that error rates are due to a "single source of variability", which is not generally the case with biometrics. Ten enrolment-test sample pairs from each of a hundred people is not statistically equivalent to a single enrolment-test sample pair from each of a thousand people, and will not deliver the same level of certainty in the results.

Effectively, the use of either the "Rule of 3" or the "Rule of 30" requires the assumption that the decisions used to estimate error rates are uncorrelated. More recently, Schuckers [9] provided a method for dealing with the issue of the dual sources of variability and the resulting correlations that arise from this structure.

The calculation given below is for the determination of the number of comparison pairs, n, from which samples need to be taken. Define a comparison pair, similar to the *enrolment-test sample pair* of Mansfield and Wayman [5], as a pair of possibly identical individuals from whom biometric data or images have been taken and compared. If the two individuals are the same then call the comparison pair a genuine one. If the two individuals are distinct then call the comparison pair an imposter one. In order to use this information to determine test size, it is necessary to specify some estimates of the process parameters before the data collection is complete. In order to obtain sample size calculations it is necessary to make these specifications. It is worthwhile noting here that most other biological and medical disciplines use such calculations on a regular basis and the U.S. Food and Drug Administration requires them for clinical trials. Approaches to carrying this out are discussed below.

Let the error rate of interest, either FMR or FNMR, for a process be represented by γ and let Y_{ij} represent the decision for the jth pair of captures collected on the ith comparison pair, where n is the number of comparison pairs, $i = 1, \ldots, n$ and j 1, …, m_i. Thus, the number of decisions that are made for the ith comparison pair is m_i, and n is the number of different comparison pairs being compared. Define

$$Y_{ij} = \begin{cases} 1 & \text{if } j\text{th decision from comparison pair } i \text{ is incorrect} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Assume for the Y_{ij}'s that $E[Y_{ij}] = \gamma$ and $V[Y_{ij}] = \gamma(1-\gamma)$ where $E[X]$ and $V[X]$ represent the mean and variance of X, respectively. Estimation of γ is done separately for FNMR and FMR and so there is a seperate collection of Y_{ij}'s for each. The form of the variance is a result of each decision being binary. The correlation structure for the $Y_{ij}'s$ is

$$Corr(Y_{ij}, Y_{i'j'}) = \begin{cases} 1 \text{ if } & i = i',\ j = j' \\ \rho \text{ if } & i = i',\ j \neq j' \\ 0 & otherwise \end{cases} \quad (2)$$

This correlation structure is based upon the idea that there will only be correlations between decisions made on the comparison pair but not between decisions made on different comparison pairs. Thus, conditional upon the error rate, there is no correlation between decisions on the ith comparison pair and decisions on

the i'th comparison pair, when $i \neq i'$. The degree of correlation is summarized by ρ. This is not the typical Pearson's correlation coefficient, rather it is the intra-class correlation or here the intra-comparison pair correlation. More details can be found in Schuckers [10].

Derivation of sample size calculations requires an understanding of sampling variability in the estimated error rate. Thus consider

$$\hat{V}[\hat{\gamma}] = N^{-2}\hat{\gamma}(1-\hat{\gamma})\left[N + \hat{\rho}\sum_{i=1}^{n} m_i(m_i-1)\right], \quad (3)$$

where $N = \sum_{i=1}^{n} m_i$, and $\hat{\gamma} = N^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i} Y_{ij}$. Fleiss et al. [11] has suggested the following moment-based estimator for ρ

$$\hat{\rho} = \left(\hat{\gamma}(1-\hat{\gamma})\sum_{i=1}^{n} m_i(m_i-1)\right)^{-1} \sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{\substack{j'=1 \\ j'\neq j}}^{m_i}(Y_{ij}-\hat{\gamma})(Y_{ij'}-\hat{\gamma}).$$

Since $\hat{\gamma}$ is a linear combination, if n is large it is reasonable to assume that the central limit theorem holds, Serfling [12]. To produce a $(1-\alpha) \times 100\%$ confidence interval for γ use

$$\hat{\gamma} \pm z_{\alpha/2}\sqrt{N^{-2}\hat{\gamma}(1-\hat{\gamma})\left[N + \hat{\rho}\sum_{i=1}^{n} m_i(m_i-1)\right]}, \quad (5)$$

where $z_{\alpha/2}$ represents the $1-\alpha/2$th percentile of a Gaussian distribution with mean 0 and variance 1. Further, if $m_i = m$ for all i (3) simplifies to

$$V[\hat{\gamma}] = (nm)^{-1}\hat{\gamma}(1-\hat{\gamma})[1+\rho(m-1)], \quad (6)$$

where N has been replaced by nm. This form will be used to derive sample size calculations.

Turning from variance estimation to sample size calculations, set the portion of (6) after the $\pm$, the margin of error, equal to some desired value B and solve for n, the number of comparison pairs. Then the following sample size calculation for making a 100 $(1-\alpha)\%$ CI with a specified margin of error of B is obtained.

$$n = \left\lceil \frac{z^2_{1-\frac{\alpha}{2}}\gamma(1-\gamma)(1+(m-1)\rho)}{mB^2} \right\rceil, \quad (7)$$

where $\lceil\ \rceil$ is the next largest integar or ceiling function. In order to create sample size calculations for a ▶ confidence interval, it is necessary to specify, among other things, the desired margin of error, B, for the interval. As mentioned above there are effectively two sample sizes when dealing with performance evaluation for biometric authentication devices. This derivation here is for the number of comparison pairs, n, that need to be tested and assume that the number of decisions per individual is fixed and known. This is equivalent to assuming that $m_i = m$ for all i and that m is known. In practice it will be possible to determine different values for n by varying m before proceeding with a evaluation. As with all sample size calculations it is important to note that specification of a priori values for the parameters in the model is necessary. In this case that means it is necessary to estimate values for γ and ρ to be able to determine the number of individuals, n. Several strategies are reasonable and have been discussed in the statistics literature for these a priori specifications. See, e.g., Lohr [13]. Ideally, it would be possible to make a pilot study of the process under consideration and use actual process data to estimate these quantities. Alternatively, it may be possible to use estimates from other studies perhaps done under similar circumstances or with similar devices. The last possibility is to approximate based upon prior knowledge without data. Regardless of the method used it is important to recognize that n is a function of α, B, m, γ and ρ. n varies directly with γ and ρ and inversely with α, m and B. Thus, a conservative approach to estimation of these quantities would overestimate γ and ρ and underestimate m. This will produce a value for n that is likely to be larger than required. Table 1 illustrates the use of (7). It is also worth noting that most studies of this type have a significant drop out rate of individuals as the data collection progresses. Thus it is adviseable to plan

Test Sample and Size. Table 1 Illustration of the use of (7)

α	B	γ	m	ρ	n
0.05	0.005	0.01	10	0.4	700
0.05	0.01	0.01	10	0.4	175
0.01	0.005	0.01	10	0.4	1,209
0.05	0.005	0.02	10	0.4	1,386
0.05	0.005	0.01	5	0.4	792
0.05	0.005	0.01	10	0.1	290

a collection process that assumes some attrition in the number of comparison pairs to be selected. The values of α and B are likely to be set by investigators or by standards bodies rather than the performance of the process under study.

Equation (7) is straightforward for calculation of the number of comparison pairs that need to be tested when γ = FNMR. It is less so when interest centers on γ = FMR. This is because for FNMR the number of comparison pairs translates to the number of individuals, while for FMR the number of comparison pairs is not proportional to the number of individuals. If all cross-comparisons are used to estimate FMR, then one can replace n with $n^*(n^* - 1)$ in (7). In that case n^* will be the number of individuals that need to be tested.

Sample Selection

Once the number of individuals to be selected is determined, another important step is to specify the target population of individuals to whom statistical inference will be made. Having done so, a sample would ideally be drawn from that group. However, this is not possible often. The next course of action is to specify a sample that is as demographically similar to the target population as possible. The group of individuals that will compose the sample is often referred to as the ▶ "volunteer crew" or simply the "sample crew", Mansfield and Wayman [5]. The more similar the sample crew is to the target population the more probable it will be that the estimates based upon the sample crew will be applicable to the target population. Often the sample crew is chosen to be a convenience sample, Hahn and Meeker [1]. Methodology for best selecting the sample crew is an open area of research in biometrics.

One useful tool for extrapolation from estimates based upon the "crew" is post-stratification. ▶ Post-stratification is a statistical tool for weighting a sample representation after the sample has been taken so that resulting estimates reflect the known population. Suppose that, there are H non-overlapping demographic groups of interest, or strata, and n_h individuals have been sampled from among the N_h total individuals in each strata. Further suppose that estimates of the error rate, $\hat{\gamma}_h$, from each of the strata are known. Then a poststratified estimate of the error rate is

$$\hat{\gamma}_{ps} = \sum_{h=1}^{H} \frac{n_h}{N_h} \hat{\gamma}_h. \tag{8}$$

An estimate of the variability of the predicted error rate is

$$\hat{V}[\hat{\gamma}_{ps}] = \sum_{h=1}^{H} \left(\frac{n_h}{N_h}\right)^2 \hat{V}[\hat{\gamma}_h], \tag{9}$$

where $\hat{V}[\hat{\gamma}_h]$ can be calculated using the equation found above. A $(1-\alpha) \times 100\%$ *poststratification* confidence interval for the process error rate can then be made using

$$\hat{\gamma}_{ps} \pm z_{\alpha/2} \sqrt{\hat{V}[\hat{\gamma}_{ps}]}. \tag{10}$$

As above, use of the Gaussian distribution here is justified by the fact that the estimated error rate, $\hat{\gamma}_{ps}$, is a linear combination of random variables.

Summary

Testing and evaluation of biometric devices is a difficult undertaking. Two crucial elements of this process are the selection of the number of individuals from whom to collect data and the selection of those individuals. Determining the number of individuals to test can be calculated based on (7). To obtain the number of individuals that need to be tested, some process quantities need to be specified. These specification can be based on previous studies, pilot studies or on qualified approximations. Selection of the "crew" for a study is a difficult process. Ideally a sample from the target population is the best, but a demographically similar "crew" is often more attainable. The inference from a demographically similar crew can be improved by the use of poststratification.

Related Entries

▶ Influential Factors to Performance
▶ Performance Evaluation, Overview
▶ Performance Measures
▶ Performance Testing Methodology Standardization

References

1. Hahn, G.J., Meeker, W.Q.: Statistical Intervals: A Guide for Practioners. Wiley, New York (1991)
2. Deming, W.E.: On probability as a basis for action. Am. Statist. **29**(4),146–152 (1975)
3. Wayman, J.L.: Confidence interval and test size estimation for biometric data. In: National Biometrics Test Center, Collected Works 1997–2000. http://www.engr.sjsu.edu/biometrics/nbtccw.pdf, pp. 89–99 (2000)
4. Wayman, J.L.: Confidence interval and test size estimation for biometric data. In: Proceedings of IEEE AutoID '99, pp. 177–184 (1999)
5. Mansfield, T., Wayman, J.L.: Best practices in testing and reporting performance of biometric devices on the web at http://www.cesg.gov.uk/site/ ast/biometrics/media/BestPractice.pdf (2002)
6. Louis, T.A.: Confidence intervals for a binomial parameter after observing no successes. Am. Statist. **35**(3), 154 (1981)
7. Jovanovic, B.D., Levy, P.S.: A look at the rule of three. Am. Statist. **51**(2), 137–139 (1997)
8. Doddington, G.R., Przybocki, M.A., Martin, A.F., Reynolds, D.A.: The NIST speaker recognition evaluation: overview methodology, systems, results, perspective. Speech Commun. **31** (2–3), 225–254 (2000)
9. Schuckers, M.E., Sheldon, E., Hartson, H.: When enough is enough: early stopping of biometrics error rate testing. In: Proceedings of the IEEE Workshop on Automatic Identification Advanced Technologies (AutoID) (2007)
10. Schuckers, M.E.: Estimation and sample size calculations for correlated binary error rates of biometric identification rates. In: Proceedings of the American Statistical Association: Biometrics Section [CD-ROM], Alexandria, VA, American Statistical Association (2003)
11. Fleiss, J.L., Levin, B., Paik, M.C.: Statistical Methods for Rates and Proportions. Wiley, New York (2003)
12. Serfling, R.J.: Contributions to central limit theory for dependent variables. Ann. Math. Stat. **39**(4), 1158–1175 (1968)
13. Lohr, S.L.: Sampling: Design and Analysis. Duxbury Press (1999)

Text-Dependent

Speaker recognition systems are said to be text-dependent when the same piece of text is used for the enrolment and for the subsequent authentication sessions. A major drawback of text-dependent systems lies in the replay attacks that can be performed easily with a simple recording device.

- ▶ Remote Authentication
- ▶ Speaker Recognition, Overview
- ▶ Speaker Recognition, Standardization

Text-Independent

Speaker recognition systems are said to be text-independent when there is no constraint on the text spoken by the user. Very convenient to use, such systems are although quite sensitive to replay attacks as any recording of the user's voice can be used to break into the system.

- ▶ Remote Authentication
- ▶ Speaker Recognition, Overview
- ▶ Speaker Recognition, Standardization

Text-Prompted

Such speaker recognition systems prompt the user to repeat a randomly chosen sequence of words. The system first performs speech recognition to verify the expected sequence of words, then the verification takes place. These systems achieve a good level of security by preventing replay attacks.

- ▶ Remote Authentication
- ▶ Speaker Recognition, Overview

Text-to-Speech (TTS)

- ▶ Voice Sample Synthesis

Text-to-Speech (TTS) Synthesis

TTS synthesis is the process of converting human readable text input into audible speech output. It involves several steps including text and linguistic analyses, letter-to-sound conversion for pronouncing each word in context, and, finally, speech synthesis to create the speech waveform.

- ▶ Voice Sample Synthesis

TFIR

The Total Frustrated Internal Reflection is one of the physical principles on which is based a contact-based optical sensor.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Thermal Biometrics

Thermal biometrics uses thermal imaging to support human identification. This is an alternative imaging modality which can reveal different (image) features for identification purposes, and it requires sensor systems which are more specialized than conventional video.

► Ear Biometrics

Thermogram

The heat pattern emitted from an object.

► Face Recognition, Thermal

Third Level Detail

This refers to the minute details of the ridges and deals with the shape and relative position of the pores, the edge shape of the ridges, and the alignment of individual ridges.

Third level detail can be the same, and if so, may add up to the value and significance of each individual point and the total. However, if different or absent, it does not prevent identification because it cannot be expected to reproduce the same in the latent and the inked print due to its minute detail, its three dimensional properties (of the source) and less ideal conditions during printing.

► Fingerprint Matching, Manual

Threshold Limit Value (TLV)

TLV® is a trademark of the ACGIH. TLVs® are similar in concept to maximum permissible exposures (MPE) and permissible exposure limits (PEL), though they are distinct. The ACGIH (www.acgih.org) defines TLVs as follows.

Threshold limit values (TLVs®) are determinations made by a voluntary body of independent knowledgeable individuals who represent the opinion of the scientific community that exposure at or below the level of the TLV® does not create an unreasonable risk of disease or injury.

TLVs® are not standards. They are guidelines designed for use by industrial hygienists in making decisions regarding safe levels of exposure to various physical agents found in the workplace. TLVs® are health-based values established by committees that review existing published and peer-reviewed literature in various scientific disciplines (e.g., industrial hygiene, occupational medicine, and epidemiology). Since TLVs® are based solely on health factors, there is no consideration given to economic or technical feasibility.

► Iris on the Move™
► Iris Device

Thresholding

Thresholding is a simple grayscale image segmentation method in which individual pixels are marked as "object" pixels if their value is greater than some threshold value (assuming an object to be brighter than the background) and as "background" pixels otherwise. Typically, an object pixel is given a value of "1" while a background pixel is given a value of "0."

► Hand Shape
► Image Segmentation

Throughput

The number of end users that a biometric system can process within a given time interval; for example, 1 min. Different applications have varying requirements on throughput. For instance, for employee attendance checking scenario, users might expect a throughput of at least 30 considering the rush hour situation.

► Performance Evaluation, Overview

Time and Attendance

Samir Tamer[1], Stephen J. Elliott[2]
[1]Ingersoll Rand Security Technologies, Campbell, CA, USA
[2]Purdue University, West Lafayette, IN, USA

Definition

Time & Attendance (TA) describes the field of tracking the hours that employees spend on the job, and compensating them appropriately. The time and attendance calculations are non-trivial, and include shift differentials, holiday pay, vacation and sick pay, flex time, minimum increments, time quantization as well as multiple employee roles, where the pay structure is different for each role.

Overview of Time and Attendance

Time & Attendance (TA) describes the field of tracking the hours that employees spend on the job, and compensating them appropriately. While such a succinct definition implies that the task is trivially simple, it is indeed not. The TA industry encompasses hundreds of competing companies focusing on one or more constituent parts of the field, such as

1. Data capture
2. Payroll calculations
3. Payroll disbursement

This article discusses the following topics: complexities of payroll calculation, data capture paradigms, a sampling of modern ► TA terminals, driving factors for choosing biometric terminals over non-biometric terminals, application-specific tradeoffs, a case study, and factors confounding the adoption of biometrics for TA.

Payroll Calculations

The value of the Time & Attendance market is the ability of companies to oversee their employees and operations. Time & Attendance facilitates the compensation of employees for time worked, and provides insight into employee attendance and absence. The benefits of accurate time calculations include a correct assessment of operational costs, improved regulatory compliance, and better management. An ancillary benefit to accurate time and attendance is employee morale – according to a report by IDC payroll accuracy is one of small businesses' biggest concerns [1].

The calculation of payroll is non-trivial. Many countries have instituted laws standardizing the amount of time an employer may reasonably expect certain classes of employee to work. Any hours above and beyond this reasonable limit are considered "overtime" and must be paid at a higher hourly rate. For example, in the United States, the Fair Labor Standards Act of 1939 (FLSA) defines the standard work week as 40 h, and requires employers to pay a 50% premium for any hours worked beyond that limit.

Note that employees must be paid overtime if they were "suffered or permitted" [2] to work those hours; thus employees can sometimes *choose* to work overtime (regardless of authorization or business need) unless specifically prohibited or controlled by the employer. The Department of Labor states that "The reason (employees work overtime) is immaterial. The hours are work time and are compensable."

The issue of overtime is confounded by the fact that different labor laws cover persons employed in different countries, states, cities, etc. The Department of Labor encourages businesses to contact local Wage and Hour Division offices to understand their responsibilities [3]. Small businesses end up becoming experts in their local labor laws, and large multinational corporations become experts in the laws at each site (including work-from-home sites) around the world. The only effective alternative to becoming a labor law expert is to outsource payroll calculations to a specialized TA company.

One example of the differences seen in laws from place to place involves the classification of certain work-hours as standard or overtime. As mentioned above, the Fair Labor Standards Act defined a standard work week as 40 h. However, in the US state of California, a standard work day is defined as 8 h, so *for certain jobs* any time beyond 8 h in a single day is considered overtime.

Other examples of calculation complexities include:

1. Shift differentials (paying a premium for working nights or weekends)
2. Holiday pay (paying a premium for working on recognized holidays)
3. Vacation & sick-pay (paying even though an employee did not work)
4. Flex time (*not* paying overtime when employees shift work hours between days)
5. Minimum increments (e.g., paying workers a minimum of 4 h for showing up at the worksite, even if they are sent home after 10 min)
6. Time quantization (e.g., rounding time to the nearest 15 min vs. the nearest minute)
7. Employees with multiple roles, who are paid different rates for each role

While this is by no means a comprehensive list, is does illustrate the point that calculating employee compensation can be very complex. It's easy to see why some companies might want to outsource this activity, freeing up time to concentrate on their core business.

Data Capture

Regardless of who calculates employee compensation, somehow the raw data must be collected. When collecting data, companies have choices on how they decide to automate and integrate. An immature time and attendance system is paper-based, and relies on the error-prone step of deciphering employees' handwriting to record their hours-worked. The next step up from this approach is a mechanical clock using time cards, as shown in Fig. 1. An employee using a mechanical clock would typically:

1. Arrive at work
2. Pick his/her card out of a rack

Time and Attendance. Figure 1 Typical punchcard based time clock.

3. Place it into a time clock to stamp the current date/time onto the card
4. Place the card back in the rack

Someone from the accounting department would then typically collect all the time cards every week and type the resulting numbers into a spreadsheet, calculate each employee's hours worked, and calculate their compensation. There is an interface with payroll, but such an interface requires manual workarounds.

As described in the previous section, there are many different rules and complexities for time and attendance systems. This complexity is sometimes solved by the third level of automation where automation of the time and attendance system encompasses the majority of workers (but not necessary all), and is automatically fed into the payroll system. It is only when full electronic time capture for all and a fully integrated system are installed that workforce management and visibility of the operation can be revealed. These various complexities in automation and integration drive the selection of a Time & Attendance system.

Today, most small businesses are semi-automated, and rely on employee vigilance to ensure that their hours were recorded correctly. These businesses also rely on the ability of the accounting department to create workarounds when errors are discovered.

Electronic data capture is inherently different, in that the ► time clock generates electronic data whenever an employee clocks-in or clocks-out. TA terminals using electronic data capture record a ► clock-in or ► clock-out transaction when the employee presents an authorized credential to the clock. The credential may be an ID number, proximity card, smartcard, magnetic stripe card, etc.

Typical Time and Attendance Terminals

There are three main classes of electronic time & attendance terminals – card or proximity based, fingerprint, and hand geometry.

Figure 2 shows a proximity card system for small business (it supports up to 250 employees). The time clock automatically calculates the total hours worked, including overtime. Employees interact with the system by bringing their proximity badge close to the reader, and a large display and internal speaker provide feedback to the employee. The system interfaces with software on Windows™ compatible machines.

There are also biometric systems for small businesses – one example is the WaspTime Biometric Time & Attendance system (Fig. 3) which uses fingerprints to recognize the employee (it supports up to 1,500 employees). The system is very similar in that it also provides clear audible and visual feedback to employees, and interfaces to software on Windows™ compatible machines. However, the larger display and numeric keypad support richer interactions with the employee.

Time and Attendance. **Figure 2** Wasp Barcode RFID reader.

Another biometric modality that is used for time and attendance is hand geometry. Figure 4 shows the HandPunch 4000 from Ingersoll Rand. Similar to the fingerprint terminal above, it supports rich interactions to thousands of employees (up to 3,498) with a clear display, audible feedback, and a keypad. The HandPunch also includes function keys and can apply validation tables to complex data entered at the terminal. Such data includes department transfers, tips collected, job codes, and pay codes [4].

Figure 5 shows the GT-400 timeclock from Ingersoll Rand, which exemplifies a recent trend in biometric time clocks: hardware manufacturers are

Time and Attendance. **Figure 3** Wasp Biometric Time and Attendance System.

Time and Attendance. **Figure 4** HandPunch 4000 from Ingersoll Rand.

Time and Attendance. Figure 5 GT-400 from Ingersoll Rand.

providing programmable Linux-based terminals so that specialized TA software companies can run their applications directly on the data-collection terminal (not only on a PC). Such terminals typically have large displays, graphical user interfaces, ATM-style keypads, and high processing power (one variant of the GT-400 claims to support >100,000 users on the terminal [5]).

Driving Factors for Choosing Biometrics

As demonstrated in the previous section, there are many commonalities between biometric and nonbiometric terminals. In fact, the only true differentiator is that biometric terminals are, well... *biometric*. Nonbiometric terminals require a physical credential (such as a magnetic stripe card or smart card) or a nonphysical credential (such as a PIN, employee number, or password) to identify the employee. However, such terminals cannot determine if the credential was presented by the authorized user or someone else. This leads to the primary driving factor for companies' choosing to employ biometric Time & Attendance systems.

Payroll fraud, also called "buddy punching," occurs when an employee's timecard is punched by someone else in order to credit that employee with more hours than were actually worked. "Buddy punching" typically involves two employees that share identification mechanisms such as tokens, passwords, or smart cards between them in order to receive pay for, or portion of pay for that day. An American Payroll Association study discussed in its January 2002 issue of PayTech Magazine found that over 5% of payroll costs in the United States are fraudulent. For many companies, shrinking or eliminating this fraud saves more money than is spent purchasing a biometric terminal.

Is buddy punching rampant in all companies? Absolutely not; it is carried out by a minority of employees, and only at some companies. Sites that are most prone to buddy punching encompass unsupervised employees and those with little management oversight. A lack of oversight can occur when any of the following conditions arise:

1. Employees work off-site or without supervision
2. Employees have flexible or staggered work-schedules
3. Employees who rise into supervisory roles feel uncomfortable disciplining their former peers
4. The workforce has high turnover, making it difficult for supervisors to know who should be where when
5. Supervisors are not present when employees clock-in or clock-out

Even sites exhibiting none of these conditions often choose to employ biometrics as an inexpensive, convenient alternative to card-based systems. Card readers generally cost less than biometric terminals, but the up-front costs of purchasing cards plus the annual cost of replacing cards or worn-out readers can tip the balance towards biometric-only or PIN + biometric systems. The convenience of biometric terminals stems from the fact that employees cannot lose or forget their biometric credential, where they can easily forget or misplace card-based credentials. When that happens, time is wasted while they track down a supervisor, spending their own time and the supervisor's time entering manual overrides to clock-in and clock-out.

Application Specific Tradeoffs

The Time & Attendance market is different than other markets. It has different customers, users, use profiles,

and product demands than other markets. For this reason, the data collection terminals focused on the TA market have different attributes than similar terminals targeting other markets (e.g., Access Control). Some of the most striking differences concern:

1. *Location* – TA terminals are generally placed in employee-only areas inside a company. Since customers and investors generally don't see these areas, there is less need for the terminal to be beautiful or svelte.
2. *User interface* – Employees sometimes verify their hours-worked or enter/extract other data from TA terminals, requiring more buttons and larger screens than AC terminals.
3. *Vandal resistance* – TA terminals must (surprisingly) be more vandal-resistant than other terminals. This is unintuitive since Access Control terminals are often located on the outside of a building where random people can attack the device with impunity. However, random people rarely bother to attack a small box hanging on the wall. Alternatively, employees who believe that a newly-installed TA system is intrusive or onerous are sometimes irritated to the point of physically damaging the system. Some even use such damage to "prove" that the system is inherently unsound and should be removed. Screwdrivers, wire snips, pliers, paper clips, and even pens/pencils can damage some TA terminals. Gum, dirt, glue, and scratches effectively disable others.
4. *Habituation* – Employees typically clock-in and clock-out every day. They grow accustomed to using the device over time, developing habits in the way they use it. This process is called "habituation" and groups of users who have gone through it are called "habituated" users. TA results in a highly-habituated workforce.
5. *Demographics* – TA workforces are often demographically diverse, including employees of differing gender, age, ethnicity, size, and job function (office workers vs. manual laborers).
6. *Error rates* – TA data collection terminals control employees' paychecks, and thus must work for every employee every time they attempt to use it. At best, failure to do so results in wasted time correcting the hours-worked. At worst it results in incorrect paychecks or lawsuits. For this reason the biometric tradeoff for TA systems is that the Failure to Enroll Rate and False Reject Rate are every bit as important as the False Accept Rate.

When assessing biometric reject rates, organizations must insist on credible (independent) data collected from a statistically-significant number of *habituated* users with an appropriate *demographic* distribution including the mix of *office workers and manual laborers* expected at the target site.

Case Study: McDonalds

This case study, excerpted from Ingersoll Rand's web site [6], is indicative of the types of applications served by biometric time clocks.

In Venezuela, McDonald's restaurants are cutting payroll costs by up to 22% annually after incorporating HandPunch biometric terminals to record time and attendance.

Over 3,400 employees at 85 McDonald's restaurants in Venezuela have been enrolled with the HandPunch over the past four years. On average, the system generates over 7,500 transactions each day resulting in over 2.5 million "punches" annually.

Students make up about 90% of the McDonald's workforce in Venezuela. They were frequently punching one another into cover for exams or other school-related events. McDonald's needed to move to biometrics to verify that the employee clocking in was really that person.

Most supervisors at McDonald's are promoted from within and many find it difficult to impose rules and restrictions on their fellow workers. The HandPunch ensures that everyone is treated fairly (Fig. 6).

Confounding Issues

Since there are so many benefits to using biometric TA terminals, why don't all companies use them? Because they have down-sides to them too. The most common issues cited when arguing against biometrics are:

1. *Morale* – Some employees find it demeaning that the employer doesn't trust them to ▶ punch-in and ▶ punch-out correctly. If some workers/departments are required to use the TA terminal while others are not (possibly because they are exempt from FLSA), tensions can rise. One employee that was dissatisfied stated "These systems are being used. . . to reduce us to mere cogs in the machine". . . "It is as if we are working in some textile mill in the

Time and Attendance. Figure 6 Biometric time clock at a quick-serve restaurant.

1820s and when the whistle blows we had better all be sitting in front of our looms ready to go... As civil servants, we produce incredibly creative, thoughtful and substantive projects but this only happens when we are treated with respect and dignity, and these systems treat us with neither!" [7].

2. *Privacy* – The idea that "big brother" is watching is unacceptable to some employees. They may link the idea of biometric TA terminals to government biometric programs they've heard about, including criminal forensics, border control, and the search for terrorists. Movies such as Minority Report invoke the image of uncontrolled biometric surveillance tracking one's every move [8].
3. *Safety* – Employees and unions often cite safety as an argument against biometric time clocks. Common issues include "lasers" in the eye (a misinterpretation of the infrared LEDs used in many face and iris systems) and microbes on the surfaces of fingerprint and hand geometry systems.
4. *Restrictive work policies* – Some TA systems incorporate rules & policies that employees find onerous, such as requiring that employees punch in/out within 5 min of their scheduled start/finish times. *This has nothing to do with biometrics*, but is often blamed on the biometric terminal since it is the most visible part of the system.

Sometimes complaints turn into grievances, and some grievances must be solved through litigation or arbitration. In the arbitration ruling of *Canada Safeway Ltd. and United Food and Commercial Workers Union, Local 401*, the arbitrator balanced the privacy rights of employees to the business needs of the company. He stated: "the more intrusive the impact on employee privacy the greater the business rationale that must be demonstrated. Conversely, if the intrusion on employee privacy is insubstantial, the concomitant level of justification also is lower." He found that Hand Geometry technology was minimally intrusive, and ruled that employees could be required to use hand geometry time clocks [9].

In another arbitration ruling, *IKO Industries Ltd. and U.S.W.A., Loc. 8580*, the arbitrator found that capturing fingerprints was not an "egregious disregard of privacy rights" but was nonetheless an invasion of privacy. IKO Industries was not able to show enough business benefit to justify that invasion, and the arbitrator ordered IKO to cease and desist their deployment of biometrics.

In addition to navigating the minefield of installing a biometric TA system, employers must be mindful of providing due process to employees during disciplinary proceedings (should they become necessary). In a non-biometric TA application in Tucson, Arizona, city sanitation workers and union representatives said that buddy-punching was allowed by supervisors, and was something that had gone on for years in the city's Environmental Services Department [10]. The firing of two supervisors was overturned because the supervisors weren't given enough warnings, and both were reinstated to their former positions and were awarded back-pay.

Companies wishing to deploy biometrics to combat buddy-punching must consider all of these confounding issues and precedents when choosing a system. Engaging with unions and educating employees *before* introducing biometric terminals can allay concerns before they escalate into grievances.

Summary

Time and Attendance is a complex field covering payroll calculations, labor law, and employee relations. Historically, the data for payroll calculations has been gathered by hand-writing time cards or using "punch"

clocks that imprint the time & date onto a time card. To decrease data entry errors and administrative overhead, companies have migrated to electronic TA systems that use an ID card or ID number to identify each employee. However, card-based and ID-based systems leave open the loophole of buddy-punching, where employees punch-in or punch-out for one another to be paid for hours they did not work. This loophole is being closed by biometric TA systems that ensure all employees are paid for exactly the hours they are physically at work.

References

1. Rowan L.: Workforce Management for SMB's: Gaining Control of Your Most Expensive Resource, IDC, February 2007
2. http://www.dol.gov/esa/whd/regs/compliance/whdfs22.pdf
3. http://www.dol.gov/compliance/guide/minwage.htm
4. http://recognitionsystems.ingersollrand.com/files/600.pdf
5. http://www.timeandtech.com/Docs/ttG-tt3-Datasheet.pdf
6. Rand, I.: Case Study #51 – HandPunch. . .McDonald's is Lovin' It!, February 16th, 2007. http://recognitionsystems.ingersollrand.com/cs/cs.php?id=35
7. Simmons, N.J.: GeoSlavery: Big Brother. Black Star News, September 5th, 2007. http://www.blackstarnews.com/?c=125&pa=3660
8. Dir. Steven Spielberg.: Minority Report. Based on short story by Philip K. Dick. Perf. Tom Cruise. Twentieth Century Fox and Dreamworks SKG, 2002
9. Clarke, G.T.: Biometrics in the workplace – where to draw the line? Personal Information Protection Act Conference 2006, April 26 and 27, 2006. Calgary, AB. http://www.governmentevents.ca/pipa2006/presentations/a1_gary_clarke_paper.pdf
10. O'Dell, R.: Buddy-punching of timecards is OK, union backers tell council. Arizona Daily Star (Tucson, AZ), September 7, 2006. http://www.azstarnet.com/metro/145557.php

Time and Attendance Terminal

A time and attendance terminal is a device that is more sophisticated than a time clock. It uses electronic data capture record to record a clock-in or clock-out transaction when the employee presents an authorized credential to the clock. The credential may be an ID number, proximity card, smartcard, biometrics etc. There are three main time and attendance terminals, card or proximity-based, fingerprint, and hand geometry.

▶ Time and Attendance

Time Clock

Time clock is a mechanical or electronic device that is used to track employee hours. When interacting with the device, employees either "punch-in" or "clock-in" when registering the start of the time, and "punch-out" or "clock-out" when leaving work or the assigned task.

▶ Time and attendance

Time Series

A time series is a sequence of data points, measured typically at successive times, spaced at (often uniform) time intervals. Learning on time series data attempts to understand the relationship between data and time, e.g., making forecasts. The order of the data points along the time axis is an important factor to be considered in the learning methods.

▶ Incremental Learning

Tongue-Print Recognition

A biometric technology for automatically identifying or verifying a person using information of tongues. As the tongue can be protruded from the body for inspection, the shape and texture information can be acquired from its images as "tongue-print" for the recognition process. Unlike face and fingerprint, the tongue is an internal organ and well protected in the mouth, so it is basically immune to forgery. This is advantageous for protecting user's privacy and security.

▶ Biometrics, Overview

Tooth Biometrics

▶ Dental Biometrics

Top and Secondary Choices

Every biometric matcher employed for user recognition compares the presented biometric with the stored templates and generates the matching scores corresponding to each of the possible user identities. The top choice refers to the user identity corresponding to user template generating the highest matching score. The secondary choices refer to the remaining choices of the possible user identities corresponding to the templates that do not generate the highest matching score. If the difference between the highest matching score and the second highest matching score is large, there is high probability that the top choice represents the correct user identity. However, if this difference is small, the top choice may not represent the correct user identity and secondary choices become important in generating the decisions.

▶ Fusion, Rank-Level

Total Transaction Time

▶ Operational Times

Touch Tablet

▶ Digitizing Tablet

Touch-Screen

A touch-screen is an electronic display that locates and captures the contact of objects within the display area. Touch-screens can be divided in two main types: those oriented to finger-input and the ones oriented to styles. Touch-screens of stylus oriented hand-held devices are based on a resistive principle and are not oriented to operation with the user fingertips. Two separated metallic layers are connected when the screen is pressed. The position of the contact point can be accurately detected, but only when the stylus is in contact with the surface, contrary to pen tablets.

▶ Signature Databases and Evaluation

Tracing

The following of adjacent parallel ridges over a certain length in the vicinity of the supposed event and by calculating whether the count of ridges in between increases or decreases, thus defining the beginning or ending of a ridge in between.

▶ Fingerprint Matching, Manual

Training

A process used to determine the values of the classifier parameters with the help of the prototypes of the data to be recognized. The data used during this step constitute the training set.

▶ Support Vector Machine

Training Data, Sufficiency

A training data set is sufficient for a learning task if all the knowledge required for correct future predictions is

contained in the data set. In practice, it is hard to judge whether a training set is sufficient for a given task. Generally, if the distribution estimated from the training data covers all possible examples that may appear in the task, then the training data is sufficient for the task.

► Incremental Learning

Training Signature

Signature used in the enrollment phase.

The training signatures are used for enrollment of a reference template, reference model parameter estimation, decision threshold estimation, and the like.

► Signature Recognition

Transfer Learning

► Incremental Learning

Transformation

The transformation refers to the process of normalizing the output (score) for a matcher to a desired range. The range of output matching scores generated from the different biometric matchers can vary significantly. This variation can be attributed to the different distance criteria used to generate matching scores or the different biometric features employed by different matchers.

► Fusion, Rank-Level

Transmission-Based Touchless Finger Imaging (TTFI)

TTFI refers to touchless fingerprint sensor.

► Fingerprint, Palmprint, Handprint and Soleprint Sensor

Transportable Asset Protection

Anthony P. Russo
Atrua Technologies, Inc, New York, NY, USA

Synonym

Asset protection

Definition

Transportable Asset Protection is a means by which personal user secrets and privileges, stored in digital form on a portable device such as a smart card or a mobile handset, are secured from unauthorized access and/or use. The assets, being carried on one's person, are highly prone to theft or loss, making the need for security that much greater. The unique challenge posed by Transportable Asset Protection is to provide adequate security and performance using the lightweight processing resources available on the mobile device in a portable and interoperable way.

Introduction

The nature of a transportable asset is that it is carried from place to place and therefore prone to theft or loss. Once lost or stolen, the vessel (or vault) holding the asset is subject to relentless physical attack by a hacker, possibly for an indefinite period of time. As our society grows more mobile, our transported assets become more valuable and our thieves more sophisticated, hence the need for greater security and privacy is increasing rapidly. Transportable asset protection in the modern sense was ushered in with the invention of smart cards. Incorporating biometric authentication into such systems is a relatively new application and research areas designed to address the unique needs of mobile users for whom the standard methods of authentication are either too cumbersome or not secure enough.

A transportable asset protection system that uses biometrics will incorporate at least one mobile computing device or host capable of storing the asset. Examples of mobile devices include smart cards, SIM cards, PDAs, mobile handsets, FLASH storage tokens, notebook computers and other such relatively

small lightweight items that can store data and/or access services. Except for the notebook computer, these devices are all examples of embedded systems, and therefore include some limited – typically *very* limited – dynamic memory and compute power.

An asset can be any digital item of value to a user. Asset types fall into three broad, somewhat overlapping categories: (1) user secrets, (2) physical or logical access to places or services and (3) rights. Secret data is most commonly in the form of username and password data that can be used to access large amounts of confidential information, such as corporate databases or encrypted documents. In more secure systems the secret may be the private key of a digital certificate, used to cryptographically encode and sign data. Accordingly, these types of assets have a hard-to-measure value to all involved parties that depends on the confidentiality of the secret.

Examples of privileged access include access to wireless or corporate networks, use of services such as internet access, online banking, or mobile-commerce transactions. While passwords and keys are sometimes used to access these services, many such systems require a secure hardware token (e.g., smart card) for multi-factor authentication. These kinds of assets usually have a directly measurable monetary value to the user and/or the service provider.

Finally, the last category of assets is the right to use a device or a digital file. Devices with restricted use are myriad, including automobiles, photocopy machines, cell phones and firearms. However, digital rights are also becoming increasingly important: the right to listen to a downloaded song or watch a movie or open a document. Digital Rights Management and transportable asset protection are likely to become increasingly coupled in the future. In most of these cases the user has a reason to prevent third parties from obtaining the asset; in other cases it is the service provider that the vested interest.

From an implementation perspective, of particular interest – due both to their importance and popularity – are those systems composed of a host (e.g., mobile handset) and a hardware token or other ▶ secure element (SE) used to store secrets and/or provide access to services. The biometric sensor is attached to the host, which usually offers little or no physical protection from hackers and may provide dubious overall security. In contrast, the secure element – physically or wirelessly connected to the host – is typically a closed, tamper-resistant, well-trusted, standardized and usually certified system with a very lightweight microcontroller and some FLASH memory or EEPROM storage. Smart cards and SIM cards are the most prevalent examples of SEs that are completely separate entities from the host. ▶ Near Field Communication (NFC) controller chips, used for short range mobile transactions, are an example of a device that can be permanently integrated into the host hardware itself. The Trusted Platform Module (TPM) is the secure element inside the high-end notebook computers that incorporate them.

A well-designed secure element is a single-chip entity designed to exacting specifications to achieve a "black box" nature. In the locked state, therefore, an SE is fairly impenetrable. However, the vast majority of these devices are unlocked by entering a 4-digit Personal Identification Number or PIN, which in itself is not terribly secure, since a PIN entry can be observed by onlookers and does not tie the unlocking process to a person as would biometric authentication. Furthermore, when cryptography is employed in a system, such as digitally signing a document, the expectation is for a very high level of security and trust. However, those having experience with such systems know that the weakest link is often in protecting the private key, something that most existing secure elements rely upon and that a simple PIN can provide. Without a biometric match, the system-wide security of a 256-bit cryptographic key is reduced from 2^{-256} to 10^{-4} per break-in attempt. Hence the need for biometrics is clear, not only for increased security but also convenience.

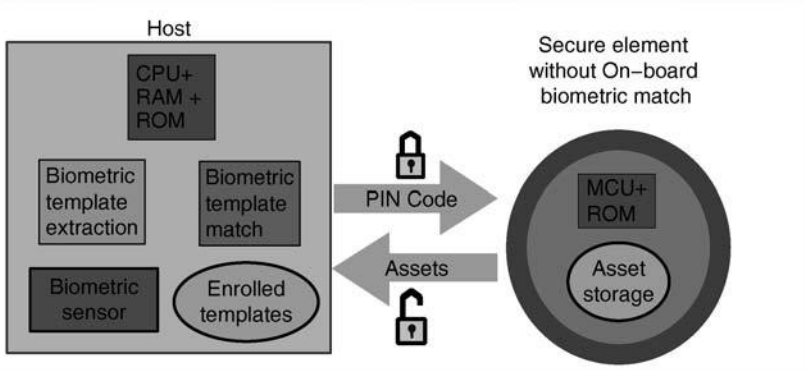
Despite the drawbacks of the PIN, in the discussion that follows the host is considered to be a non-secure entity while the secure element is maximally trusted. Therefore, the more processing done on the secure element, the safer the system. As will become clear, the most secure platforms are the ones in which the host and the secure elements are the same physical device.

System Architectures

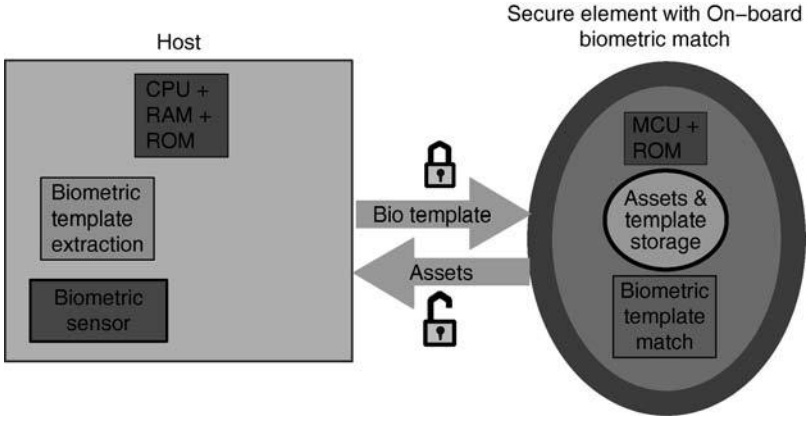
Implementations of transportable asset protection systems are varied. The most trivial, least secure architecture employs only a host, with no secure element connected to the system. Security risks abound, the

most important of which is the lack of any secure element to protect the assets. While the host itself might employ security features such as user locking, those are easily circumvented by a thief who steals the phone and accesses the filesystem through alternate hardware means. In short, this architecture is too insecure to constitute a useful asset protection system.

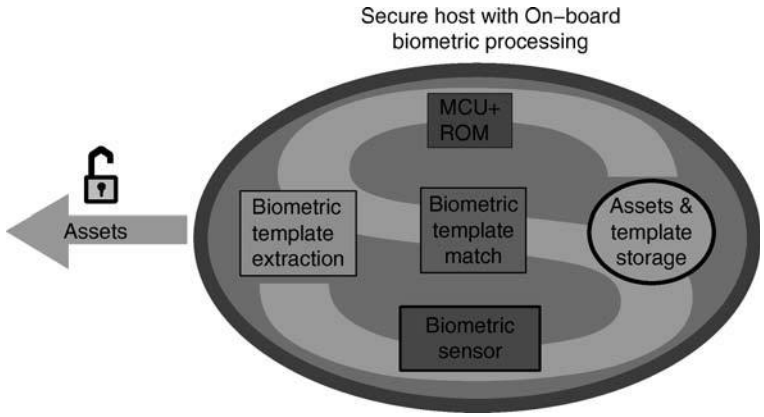
However even those implementations that use a secure element can vary widely in effectiveness. Figures 1, 2 and 3 show examples of transportable asset protection systems that offer basic security (Fig. 1), very good security (Fig. 2) and maximal security (Fig. 3). Referring to the basic security architecture, a secure element is used only to protect the assets and nothing else. In this case, the biometric sensor, template extraction algorithms, and matcher all execute from the host CPU, and enrolled template data is stored on the host's filesystem. This implementation is the easiest to engineer because typically the host has much more processing power and memory than a secure element. Assets are protected with only a PIN, but putting aside that shortcoming for the moment, a number of *additional* security risks exist with such a system.



Transportable Asset Protection. **Figure 1** Diagram describing a basic transportable asset protection system employing a host and a Secure Element. The host is connected to a biometric sensor and performs all template extraction tasks. The secure element is used only to store assets, and must be unlocked using a PIN. Once unlocked, the secrets are accessible by the host.



Transportable Asset Protection. **Figure 2** Diagram of a transportable asset protection system employing the MOSE architecture. The host is connected to a biometric sensor and performs all template extraction tasks. Assets and enrolled templates are saved on the secure element, which has the ability to perform a biometric match. The SE is unlocked via a PIN or by sending it a live scan template that matches the stored enrolled template. Once unlocked, the secrets are accessible by the host.



Transportable Asset Protection. **Figure 3** Diagram of a transportable asset protection system employing a host that is itself a secure element. All processing is performed on the secure host, making it the most trustworthy, albeit the most costly design choice.

The first security hole is that templates must be stored on the filesystem. This means these templates are readable by anyone operating the host device, representing both a security and a privacy issue, as the templates themselves could be considered an asset. While the natural inclination is to encrypt this data, the problem is that encryption requires a private key, and that key must be stored someplace for use later during decryption. Usually the key is hidden somewhere on the filesystem or in the code itself, but this kind of "security by obscurity" is a major risk nonetheless [1]. If the host does not guard against tampering, then it would be quite easy for a hacker to access the data in unencrypted form by tapping the memory bus during the match process, where it must be in plaintext form prior to use. Even easier hacks involve modifying the results of the matching algorithm to force a match even when one does not exist. [2]

A variation on this system uses the secure element to simply *store* the enrolled templates, thereby removing the need to store them on the host and offering increased protection if the host is lost or stolen. In this case, the templates are stored safely on the tamper-proof secure element, where presumably only very advanced hacking techniques could unlock it. This approach offers only a marginal increase in security, however, if the host and the secure element are both lost, which can often be the case (e.g., a mobile phone + SIM card). The memory tapping techniques are still effective as the secrets must leave the card before they are matched, and altering the match results is still possible because the match occurs on the host. Of note is that a PIN must still be used to unlock the secure element before a biometric match can be performed, and though this can be considered a security feature, it removes the user convenience that is a major selling point of biometric implementations, since PIN entry, especially on a small mobile device, is often cumbersome.

A much more secure approach is shown in Fig. 2, where the matching algorithm is executed inside the secure element itself, thereby eliminating the need for PIN entry entirely. This configuration is commonly referred to as ▶ Match-on-Card (MOC) when used with a smart card or SIM card [3]. A more general term is Match-on-Secure-Element (MOSE). In this architecture, the sensor and template extraction are still performed on the host, but during the verification process, the template is sent to a locked secure element, which only opens upon a successful match against the enrolled templates stored there. This approach has a number of important advantages. Firstly, the enrolled template data, once stored, never leaves the secure element for any reason; it is well-protected under all circumstances. Secondly, it is not possible to simply alter the match results because the matcher executes in a safe, black box environment. One other important aspect of this implementation is that it allows for secure replacement of the PIN with a biometric. This preserves both the convenience and security largely responsible for the market success of biometric systems.

The attacks on such systems tend to focus on the host processing: both sensor data and extracted templates can be snooped and modified, allowing for

replay and denial-of-service attacks. These issues are addressable for the most part, as this class of hacks can only be launched during the authentication process and therefore will be useless if the device is stolen afterward. Overall, this is the most cost-effective and technologically feasible approach to transportable asset protection.

The least cost-effective approach, but offering the highest level of security, is one in which the sensor, template extraction, and matching are *all* located on the secure element (see Fig. 3). In this case it is not possible for a hacker to snoop or alter any biometric data or algorithms. The most promising attacks on such a system involve fake-finger spoofing, for which there are countermeasures available. However, such a system is costly, because the microcontroller on the secure element can no longer be lightweight: much more memory and computing power would need to be available to process the raw biometric data into a template and perform the match. Other considerations involve powering such a secure element and addressing the mechanical mounting issues presented by placing a sensor on a smart card or other miniature device, where it must meet flexibility and height standards to be universally adopted.

Role of Standards

As noted, the secure elements used to protect transportable assets are typically standard-based tokens such as smart cards and SIM cards. Being portable, standards have evolved so that these tokens can interoperate with a variety of different host hardware, as long as the token and the host support the same standard. For commercial deployments, it has become increasingly difficult to attract interest in proprietary solutions and formats due to this expectation of interoperability.

A good example to illustrate the need for standards is a GSM-based mobile handset, which uses a SIM card to verify the user and provide wireless network access. Let's assume this is a MOC system where the biometric sensor and template extraction algorithms are on the handset, and the removable SIM card has embedded in it a biometric matcher where the enrolled templates are stored on the SIM card. Given that wireless network operators require that a SIM card be usable in any GSM handset, it is prudent to ensure that handsets adhere to the same template format standard that the SIM card does, or else the biometric matching would not work. While it is possible and sometimes necessary to fall back to having the user enter a PIN instead, this scenario is undesirable from the operators' point-of-view as it weakens the value of deploying biometric SIM cards to their network subscribers.

The ISO-7816 standard [4] is an example of how smart cards interact with their hosts at both an electrical, physical, and communication level, and ISO-7816-11 even allows for biometric Match-on-Card commands. But as the above example illustrates, standard commands are not enough to ensure compatibility. In a MOSE architecture the host and secure element must agree on the nature and format of biometric templates to be used for enrollment and unlocking. Although some widely-accepted standards exist for face recognition [5] and other biometrics [6], fingerprints currently offer the most options in this regard, and the only ones thus far designed to directly address the requirements of a MOSE system [7].

Algorithmic Challenges and Approaches

To fully appreciate the challenges involved with creating a biometric matcher for use within a secure element, it is necessary to understand the execution environment in which it runs. The most common secure elements employ 8-bit or 16-bit microcontroller units (MCUs), having available RAM from 500 bytes to 2KB, and 128KB or less of ROM. Internal clock rates are typically below 30 MHz, often in the 7–15 MHz range. There is usually no floating-point support, and sometimes no native hardware support for signed integer math, as is the case with an 8051-class MCU [8]. The computational power of the Pentium-class CPU found on most PCs is roughly 3 orders of magnitude more than that of the MPU found on most secure elements; available RAM is 6 orders of magnitude more. While the processing power for secure elements does increase from year to year, low-end security processors continue to dominate the mass market due to its low cost. Therefore, it is important to design it with these in mind.

The simplest overall approach is to design the biometric template so that it is a statistical feature vector, allowing matching to be done via a linear or non-linear

classifier, which is more suitable for implementation on low-end MCUs. Certain biometrics lend themselves more easily to lightweight matching algorithms than others. Iris matching and DNA matching are two examples where most of the computational effort is in the template extraction step, while the matching step is relatively straightforward. Standard fingerprint matching does not fall into this category – it typically requires both local and global geometric information – which is unfortunate given its popularity with and suitability for mobile platforms.

Nonetheless, most of the implementation issues have been overcome, and fingerprint-based MOSE systems are among the most popular, including Giesecke & Devrient's STARCOS 2.4 smart card Operating System with On-card matcher [9]. The initial FP-based MOSE algorithms, first introduced in 1999, used proprietary templates [10, 11] for which much of the data was preprocessed on the host to minimize the computational load on the secure element. Second-generation algorithms use a standards-based template format [7] but, due to the added burden this puts on the secure element, accuracy tradeoffs have been necessary to achieve acceptable match speed. Current work underway by the author promises a more complex third-generation, standards-based minutiae matching algorithm with much smaller accuracy degradation and suitability for execution on most secure elements.

Of note is that pattern matching techniques for fingerprints, with their relatively large template size and image-processing-based algorithms, are harder to implement and in many cases simply not practical, leaving the aforementioned statistical feature-vector approach as the best alternative to minutia-based methods. Unfortunately, in addition to accuracy issues, proprietary statistical approaches suffer from lack of any widely-accepted template standards, drastically reducing their suitability for the marketplace.

Recent Trends

The overall trend for biometrics – especially fingerprints – is rapidly growing in the consumer market, most notably in notebook computers, mobile phones, door locks and USB storage tokens. While some of these applications focus on the convenience aspect of biometrics (e.g., password replacement) almost all of them perform some sort of asset protection functionality, with varying levels of security depending on the design and application.

Similarly, MOSE systems for asset protection continue to gain in popularity, and deployments are expected to grow quite rapidly in the next 2–5 years. Mobile handset applications, which employ NFC chips and/or SIM cards, are driving the commercial need. Atrua Technologies, Inc., by which the author is employed, has demonstrated the first SIM card with on-chip fingerprint matching in 2007 as well as the first NFC controller with on-chip matching. It is expected that various vendors will offer MOSE capable systems in the near future and that mobile handsets with MOSE capabilities will reach the mass market worldwide within this timeframe.

On the government side, the U.S. has recently put together guidelines for testing Personal Identity Verification (PIV) cards [12] with fingerprint-based match-on-card capability using FIPS 201 minutiae format standards. As PIV cards have become the standard way of identifying U.S. Government employees, the added security and convenience of match-on-card will see a wide deployment if the planned technology trials are successful.

Related Entries

- ▶ Template Security
- ▶ Biometrics and Security, Standardization
- ▶ Common Biometric Exchange Formats Framework, Standardization
- ▶ Fingerprint Matching, Automatic

References

1. Schneier, B.: Applied cryptography, Wiley, Inc., New York (1996)
2. Maltoni, D., Maio D., Jain, A.K., Prabhakar, S.: Handbook of fingerprint recognition, Springer, New York (2003)
3. Russo, A.P.: "Fingerprint-based authentication and smart cards: Issues and trends", E-Payments 2000 Conference. http://www.epf.net/PrevMtngs/Sep00/Sep00Meeting.html. Accessed Oct 5, 2000
4. ISO/IEC 7816-4: "Information technology–Identification cards–Integrated circuit(s) cards with contacts, Part 4: Interindustry commands for interchange," International Standard (1995)
5. ANSI-INCITS 385-2004: *American National Standard for Information Technology – Face Recognition Format for Data Interchange* (2004)

6. ANSI-INCITS 398-2005: *American National Standard for Information Technology – Common Biometric Exchange Formats Framework (CBEFF)* (2004)
7. ISO/IEC 19794-2: *Information Technology—Biometric Data Interchange Formats—Part 2: Finger Minutiae Data* (2005)
8. MacKenzie, I.S., Phan, R.C.-W.: The 8051 Microcontroller, Prentice Hall, (2006)
9. Giesecke & Devrient: STARCOS® 2.4 Card Operating System, Bio Version http://www.gdai.com/portal/page-_pageid=42,70526&_dad=portal&_schema=PORTAL.htm. Accessed Dec 30, 2005
10. Russo, A.P.: Fingerprint matching algorithm for low-cost 8-bit smart cards. RSA Conference Europe 2000 (2000)
11. Russo, A.P. (inventor): Method and system for fingerprint template matching. U.S. Patent Number 6681034 (1999)
12. National Institute of Standards and Technology: *FIPS Pub 201-1 Personal Identity Verification (PIV) of Federal Employees and Contractors*. Federal Information Processing Standards Publication (2006)

Tread Pattern

The outsole or underside of a shoe can sometimes have a complex pattern of ridges and various shapes which provide traction for the wearer. For example, athletic shoes may have a tread pattern composed of many slanted ridges placed close together so as to give an improved grip to the running surface, while hiking shoes may have a thick and high profiled wedge-shaped tread for easy walking in soft terrain. This complex pattern is commonly referred to as a tread pattern.

► Footwear Recognition

Trusted Biometric System

Biometric system with template protection, compared to traditional biometric system, paying more attention to security and privacy issues of the users.

► User Interface, System Design

Trusted Traveler

► Registered Traveler

U

UBM

► Universal Background Models

ULW

► Universal Latent Workstation

Unauthorized Data Collection

In principle, all data collected must be backed by the law, whether state legislation or contractual obligations. A data item may well be necessary for a particular purpose, but this does not mean that the purpose itself is authorized, or that the agent (human or machine) collecting the data is empowered to do so.

► Privacy Issues

Unauthorized Data Disclosure

From a privacy perspective, the end-user owns his or her own personal data, even if the data resides in an organization's computer system. Therefore, permission should be sought from the end-user before data is disclosed or shared. In practice, different jurisdictions have different notions of data ownership, which may result in disclosure without the end-user's explicit consent or even knowledge.

► Privacy Issues

Unification Framework

A unification framework includes a collection of fusion algorithms and it uses the evidences obtained from the input biometric probe data to dynamically select the optimal fusion algorithm. The selected fusion algorithm is then used to compute the fused biometric information. The unification or reconciliation should satisfy most of the application requirements and yield better recognition performance.

► Fusion, Sensor-Level

Universal Background Models

Douglas Reynolds
MIT Lincoln Laboratory, Lexington, MA, USA

Synonyms

General model; Person-independent model; UBM; World model

Definition

A Universal Background Model (UBM) is a model used in a biometric verification system to represent general, person-independent feature characteristics to be compared against a model of person-specific feature characteristics when making an accept or reject decision. For example, in a speaker verification system, the UBM is a speaker-independent Gaussian Mixture Model (GMM) trained with speech samples from a large set of speakers to represent general speech characteristics. Using a speaker-specific GMM trained with speech samples from a particular enrolled speaker, a

© 2009 Springer Science+Business Media, LLC

likelihood-ratio test for an unknown speech sample can be formed between the match score of the speaker-specific model and the UBM. The UBM may also be used while training the speaker-specific model by acting as a the prior model in Maximum *A Posteriori* (MAP) parameter estimation.

Likelihood Ratio Test

To understand the development and use of a Universal Background Model (UBM), the likelihood-ratio test for which it is intended is first described. Given an observation, O, and a person, P, the task of verification is to determine if O was from P. This verification task can be restated as a basic ▶ hypothesis test between

$$H_0 : O \text{ is form person } P$$
$$H_0 : O \text{ is } \underline{\text{not}} \text{ form person } P$$

Using statistical ▶ pattern recognition techniques, the optimum test to decide between these two hypotheses is a likelihood ratio test (Strictly speaking, the likelihood ratio test is only optimal when the likelihood functions are known exactly. In practice this is rarely the case.) given by

$$\frac{p(O|H_0)}{p(O|H_1)} \begin{cases} \geq \theta \text{ Accept } H_0 \\ < \theta \text{ Reject } H_0 \end{cases}, \qquad (1)$$

where $p(O|H_i)$, $i = 0,1$ is the probability density function for the hypothesis H_i evaluated for the measurement Y, also referred to as the "likelihood" of the hypothesis H_i given the measurement ($p(A|B)$ is referred to as a likelihood when B is considered the independent variable in the function.). The decision threshold for accepting or rejecting H_0 is θ. The basic aim in developing a verification system is to determine techniques to compute this likelihood ratio function, usually by finding method to represent and model the two likelihoods, $p(O|H_0)$ and $p(O|H_1)$.

The first step in a verification system is to extract from the observation features that convey person-dependent information, such as vocal-tract related spectral measurement when the observations are speech samples in a speaker verification system. The output of this stage is typically a sequence of feature vectors representing the observation, $X = \{x_{1,\ldots,}x_T\}$. These feature vectors are then used to compute the likelihoods of H_0 and H_1.

In statistical pattern recognition based verification systems, H_0 is represented by a model denoted λ_P, that characterizes the distribution of features derived from observations from the person P in the feature space of x. For example, one could assume that a Gaussian mixture model (GMM) distribution best represents the distribution of feature vectors for H_0 so that λ_P would be denoting the weights, means, and covariance matrix parameters of a GMM. The alternative hypothesis (see entry on ▶ GMM for more details of this model). H_1, is likewise represented by a model $\lambda_{\overline{P}}$. The likelihood ratio statistic is then formed as

$$\mathrm{LR}(X) = \frac{p(X|\lambda_p)}{p(X|\lambda_{\overline{p}})} \qquad (2)$$

See other articles in this book and chapters in [1, 2] for more details on speaker verification systems.

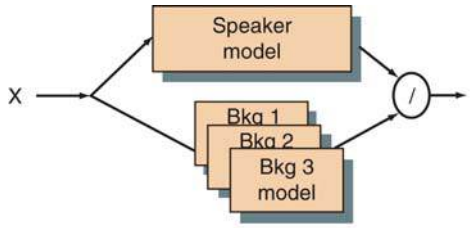
Alternative Hypothesis Modeling

While the model for H_0 is well defined and can be estimated using training samples from P, the model for $\lambda_{\overline{P}}$ is less well defined since it must potentially represent the entire space of possible alternatives to the person P.

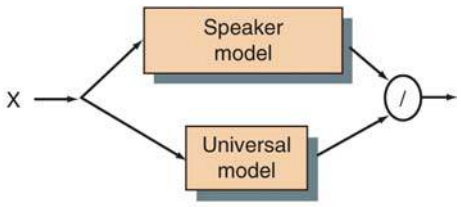
From the area of speaker recognition, two main approaches have been taken for this alternative hypothesis modeling. Here terms "speakers" and "speech samples" are used, but these apply equally to other biometric measurements and features. The first approach is to use a set of other speaker models to cover the space of the alternative hypothesis. In various contexts, this set of other speakers has been called likelihood ratio sets [3], cohorts [4] and background speakers [5]. Given a set of N background speaker models $\{\lambda_1, \ldots, \lambda_N\}$, the alternative hypothesis model is represented by

$$p(X|\lambda_{\overline{P}}) = \mathcal{F}(p(X|\lambda_1), \ldots, p(X|\lambda_N)), \qquad (3)$$

where $\mathcal{F}()$ is some function, such as average or maximum, of the likelihood values from the background speaker set. The selection, size, and combination of the background speakers have been the subject of much research (for example [4–6]). In general, it has been found that to obtain the best performance with this approach requires the use of speaker-specific background speaker sets. This can be a drawback in an



$$p(X \mid \lambda_{\overline{p}}) = f\left(p(X \mid \lambda_b, b = 1,\ldots, B)\right)$$



$$p(X \mid \lambda_{\overline{p}}) = p(X \mid \lambda_{ubm})$$

application using a large number of hypothesized speakers, each requiring its own background speaker set.

The second major approach to alternative hypothesis modeling is to pool speech from several speakers and train a single model. Various terms for this single model are a general model [7], a world model, and a universal background model [8]. Given a collection of speech samples from a large number of speakers representative of the population of speakers expected during recognition, a single model, λ_{bkg}, is trained to represent the alternative hypothesis. Research on this approach has focused on the selection and composition of the speakers and speech used to train the single model [9, 10]. The main advantage of this approach is that a single speaker-independent model can be trained once for a particular task and then used for all hypothesized speakers in that task. It is also possible to use multiple background models tailored to specific sets of speakers [10, 11].

Universal Background Models

Most modern speaker verification system use a UBM for modeling the alternative hypothesis in the likelihood ratio test. Typically, GMMs are used for distribution models and a speaker-specific model is derived by using MAP estimation with the UBM acting as the prior model (see article on GMMs for details on MAP estimation). In the GMM–UBM system a single, speaker-independent background model is used to represent $p(X|\lambda_{\overline{p}})$. The UBM is a large GMM (2,048 mixtures) trained to represent the speaker-independent distribution of features. Specifically, speech samples are selected that are reflective of the expected alternative speech to be encountered during recognition. This applies to both the type and quality of speech, as well as the composition of speakers. For example, for a verification system using telephone speech and only male speakers, the UBM would be trained using telephone speech from a pool of male speakers. In the case where such specific prior knowledge of the gender composition of the alternative speakers is not known, speech samples from both male and female speakers are used. Other than these general guidelines and experimentation, there is no objective measure to determine the right number of speakers or amount of speech to use in training a UBM.

Given the data to train a UBM, there are many approaches that can be used to obtain the final model. The simplest is to merely pool all the data and use it to train the UBM via the EM algorithm. One should be careful that the pooled data is balanced over the subpopulations within the data. For example, in using gender-independent data, one should be sure that there is a balance of male and female speech. Otherwise, the final model will be biased towards the dominant subpopulation. The same argument can be made for other subpopulations such as speech from different microphones. Another approach is to train individual UBMs over the subpopulations in the data, such as one for male and one for female speech, and then pool the subpopulation models together. This approach has the advantages that one can effectively use unbalanced data and can carefully control the composition of the final UBM. Still other approaches can be found in the literature (see for example [10, 12]).

The concept of a UBM is also used for discriminative systems, such as Support Vector Machines (SVM), where explicit likelihood functions for the two hypothesis are not used. In this case, the UBM refers to the collection data from the general population used as negative examples while training a person-specific discriminate function [13].

Acknowledgment

This work was sponsored by the Department of Defense under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

Related Entries

► Gaussian Mixture Models
► Session Effects on Speaker Modeling
► Speaker Matching
► Speaker Recognition, Overview

References

1. Benesty, J., Sondhi, M., Huang, Y. (eds.): Springer Handbook of Speech Processing, vol. XXXVI. Springer, Berlin (2008)
2. Müler, C. (ed.): Speaker Classification I: Fundamentals, Features, and Methods. Volume 4343/2007. Springer: Lecture Notes in Computer Science, Berlin (2007)
3. Higgins, A., Bahler, L., Porter, J.: Speaker verification using randomized phrase prompting. Digital Signal Process. **1**, 89–106 (1991)
4. Rosenberg, A.E., DeLong, J., Lee, C.H., Juang, B.H., Soong, F.K.: The use of cohort normalized scores for speaker verification. In: International Conference on Speech and Language Processing, Banff, Alberta, Canada, pp. 599–602 (1992)
5. Reynolds D.A.: Speaker identification and verification using Gaussian mixture speaker models. Speech Commun, **17**, 91–108 (1995)
6. Matsui, T., Furui, S.: Similarity normalization methods for speaker verification based on a posteriori probability. In: Proceedings of the ESCA Workshop on Automatic Speaker Recognition, Identification and Verification, Martigny, Switzerland, pp. 59–62 (1994)
7. Carey, M., Parris, E., Bridle, J.: A speaker verification system using alphanets. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Toronto, Canada, pp. 397–400 (1991)
8. Reynolds, D.A.: Comparison of background normalization methods for text-independent speaker verification. In: Proceedings of the European Conference on Speech Communication and Technology, Rhodes, Greece, pp. 963–967 (1997)
9. Matsui, T., Furui, S.: Likelihood normalization for speaker verification using a phoneme- and speaker-independent model. Speech Commun. **17**, 109–116 (1948)
10. Rosenberg, A.E., Parthasarathy, S.: Speaker background models for connected digit password speaker verification. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Atlanta, Georgia, USA, pp. 81–84 (1996)
11. Heck, L.P., Weintraub, M.: Handset-dependent background models for robust text-independent speaker recognition. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Munich, Germany, pp. 1071–1073 (1997)
12. Isobe, T., Takahashi, J.: Text-independent speaker verification using virtual speaker based cohort normalization. In: Proceedings of the European Conference on Speech Communication and Technology, Budapest, Hungary, pp. 987–990 (1999)
13. Campbell, W.M.: Generalized linear discriminant sequence kernels for speaker recognition. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Orlando, Florida, USA, pp. 161–164 (2002)

Universal Latent Workstation

Tom Hopper
Criminal Justice Information Services Division, FBI (Retired)

Synonym

ULW

Definition

In the past, the commercial automated fingerprint identification systems (AFIS) used by local, state, and federal criminal justice agencies have not accounted for cross-jurisdictional searching of latent crime scene fingerprints. The mobile criminal might avoid identification by simply crossing a jurisdictional boundary. This lack of interoperability stems from the fact that each AFIS vendor has a unique set of features to characterize and match fingerprints. Searching multiple AFIS in this environment involves redundant encodings of the latent fingerprint on separate workstations, each using a different vendor's feature set.

The universal latent workstation (ULW) simplifies cross jurisdictional searches by enabling an examiner to search multiple AFIS with a single fingerprint feature encoding. In many cases, the examiner will edit the features to optimize the search for a particular AFIS but they will not need to reenter the case. The ULW is based on an open interface standard developed in cooperation with the AFIS vendors,

Universal Latent Workstation. **Figure 1** Fingerprint ridge flow pattern classifications: Arch (left), Loop (center), Whorl (right).

effectively establishing a common language for crime scene investigators to share latent fingerprint identification services. The FBI provides the ULW software, training, and support to criminal justice agencies at no charge.

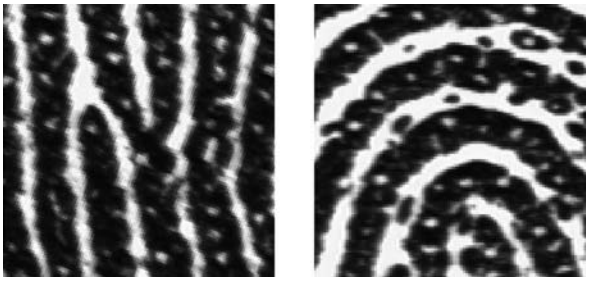
Introduction

Fingerprints have a long history in the investigation of crime and play a major part in criminal apprehensions and convictions [1]. Identifications start with the discovery, development, and capture of ▶ latent fingerprints at the crime scene or on an object used in connection with the commission of a crime. The print is then scanned into a latent workstation where the ▶ features are extracted and formatted into a search record. The search record is then sent to an AFIS to find candidate matches for the latent print. A fingerprint examiner makes the final comparison and identification.

Within the US, every state and most of the larger cities have an AFIS with arrest records from within their jurisdiction. At the national level the FBI has the integrated automated fingerprint identification system (IAFIS) with arrest prints for the entire country, currently about 55 million records. A latent print search that is not identified in the state AFIS must be searched in the national system to determine if it matches arrest prints from another state. ULW was developed to facilitate searching a single latent fingerprint in multiple AFIS.

Fingerprint Features

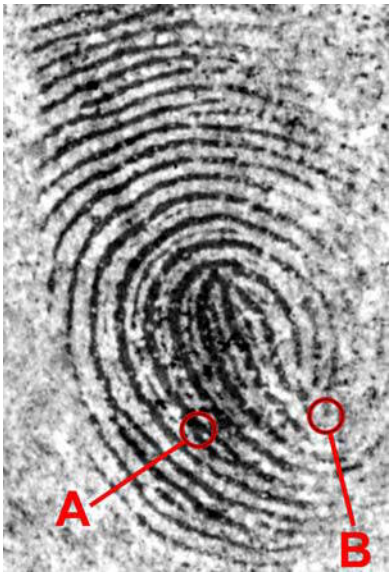
Fingerprints comprise ridges and furrows in a flow like pattern. The characteristics or information content of



Universal Latent Workstation. **Figure 2** Fingerprint features: ridge ending and bifurcation (left), pores, dots, and distinctive ridge edge detail.

the ridge structure is traditionally categorized into three levels of detail. Level 1 includes the general ridge flow and pattern classification (Fig. 1). A particular ridge flow pattern is not unique to an individual but can be used in filtering out or excluding a portion of a data base from further consideration [2]. If the crime scene print is a whorl, for example, there is no need to search arches and loops; however this strategy can occasionally cause a miss at the classification boundaries [3].

Level 2 details are the ridge path characteristics including bifurcations, endings, and dots. The normal parallel flow of the ridges is occasionally disrupted when a ridge ends or bifurcates into two ridges (Fig. 2). These disruptions, called minutia, occur at random locations and are the primary basis for identification. Level 3 features are the sweat pores and edge texture that make up the finer details along the ridges. When level 3 features are visible, they provide the examiner with additional points of comparison to reach an identification or exclusion decision [4]. Current research is bringing level 3 detail into the automated AFIS search as well [5].



Universal Latent Workstation. Figure 3 Exceptionally high clarity latent print. The bifurcation (**a**) and the ridge ending (**b**) are on the same ridge and the ridge count between them is four intervening ridges. You can use these relationships to compare this latent to the matching print in Fig. 1.

Fingers are pliable and ▶ distortions in the fingerprint may cause some shifting in the positions of the minutiae. By focusing on the topological structure of the minutiae, the examiner can accurately assess the similarity between two fingerprints without adverse effects due to variance in minutia locations. The number of intervening ridges between corresponding minutia pairs will be the same even if one of the impressions is stretched (Fig. 3). Additionally, an examiner may follow a ridge path forward from a minutia to see if another event affects the same ridge.

Search Preparation with ULW

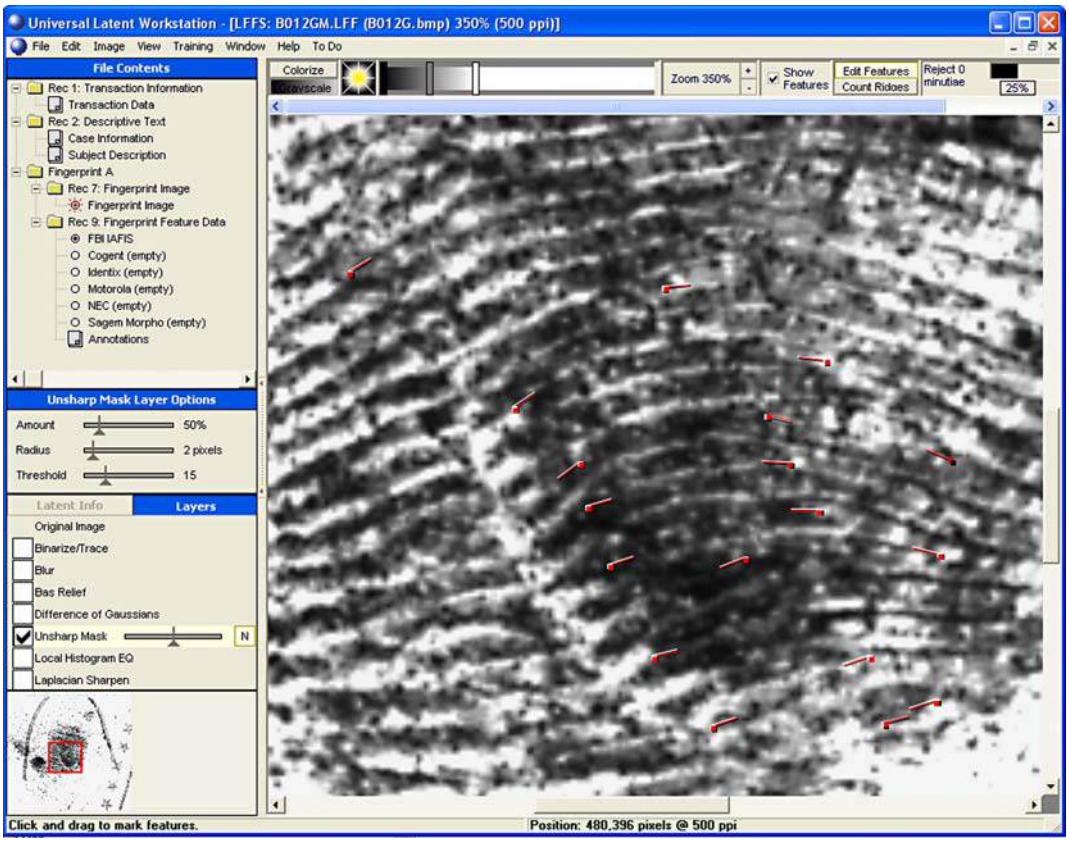
The clarity of a fingerprint image determines how difficult it will be to locate and mark the minutiae. Many latent prints are low contrast partial impressions on a substrate or surface with texture and graphics that make it difficult to follow the ridge structure. On these low quality prints, automated feature extraction may not locate all the correct minutiae and will likely mark several false minutiae. In this case, the examiner will have to mark the minutiae in the areas where the auto ▶ encoder could not follow the ridges. The ideal work environment for the examiner would be to drop the false minutiae from the auto encoding and display only the correct minutiae for editing and verification. Unfortunately, the software does not know which minutiae are false; otherwise they would not have been included in the first place. ULW addresses this problem by assigning a confidence score to each minutia based on the local image quality. ULW will then only display minutiae above a threshold confidence score and the examiner can adjust the threshold to their own preference (Fig. 4).

Counting the number of intervening ridges between neighboring minutiae has similar issues. It is a very tedious task to do manually and on a clear image the software will provide accurate results. As the image quality degrades, a procedure is needed to split the task between the software and the examiner. ULW assigns a confidence score to each ridge count based on the clarity of the ridges traversed. In ridge count verification mode, ULW will present the ridge counts in score sequence starting with the lowest. Once the examiner reviews several consecutive ridge counts without finding any errors it is likely that all the remaining ridge counts are correct as well.

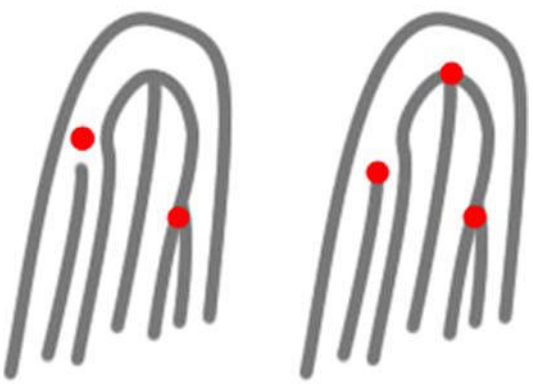
In addition to the fingerprint features the search record can include variables to limit the search to a subset of the file. Descriptors such as sex, eye color, and weight range can be entered if available. Also, the search can be focused on a geographic region by entering state codes. In practice, the variables that are most commonly used are the finger number and pattern classification. Focusing the search can improve accuracy with a large system. When searching a latent print with limited information a small percentage of the file prints will tend to get high scores in the same range as the true match. Reducing the size of the search population will reduce the high scoring non-matches and improve the chances that the true match makes it into the candidate list [6].

Searching Multiple AFIS

The commercial AFIS used each by local, state, and Federal criminal justice agencies have different rules for marking the minutiae and other features on a latent search. Figure 5 shows a diagram of two



Universal Latent Workstation. **Figure 4** ULW screen display for marking minutia. This latent print matches the arch print in Fig. 1.



Universal Latent Workstation. **Figure 5** Two different vendor encodings for the same print.

different vendor's encoding for the same print. The minutiae placement rules for vendor on the left require that the ridge ending be marked in the center of the valley just beyond the actual end of the ridge. The bifurcation in the center of the print is not marked because it is in a high curvature region. The other vendor, on the right, expects the ridge ending to be marked right on the ending of the ridge and also expects minutiae in high curvature areas to be marked. In addition to the differences in the placement of minutiae, some vendors require ridge counts to four neighboring minutia, some require ridge counts to eight neighbors and others do not use ridge counts [7]. In practice, the differences are important for accuracy but do not present a problem for the examiners.

ULW can create a latent search record for any of the major AFIS vendors or translate a record from one format to another. In some states ULW is used to search both the state system as well as the FBI IAFIS. The examiner first marks the features for the state system and if that search does not find a match, they can then translate the state search into an IAFIS search. IAFIS requires eight ridge counts per minutia so ULW would calculate the additional ridge counts. The

examiner would then verify and edit the new ridge counts. Any ridge counts from the state search that had already been verified would have a confidence score of 100 so they would not be verified a second time. Even if the state search was conducted with a latent workstation provided by the state AFIS vendor, the search record can still be imported into ULW and then converted using the same process in order to conduct an IAFIS search.

The search records are formatted in accordance with the ANSI/NIST standard: Data Format for the Interchange of Fingerprint, Facial, & Other Biometric Information [8]. This is a tagged field format and a block of fields have been reserved for each vendor. The development of the standards and its use for latent searching has been a cooperative effort by state and local agencies, the AFIS vendors, NIST, and the FBI. The ULW was first developed as a proof of concept for standards based latent searching and is now in use across the country. There are currently about 200 agencies using ULW to conduct almost 10,000 searches a month. Approximately, 10% of the cases that make it up to a National level search produce identifications.

Summary

Any effort to combat crime and the threat of terror is dependent on cooperation and sharing information across agencies. Sharing latent identification services within the US is complicated by the hierarchical network of AFIS systems and the variation in latent search feature sets. The FBI developed ULW to address these issues and facilitate information sharing across the Nation's city, county, and state borders.

Related Entries

► Biometric Standards for Law Enforcement Applications
► Fingerprint Features
► Large Scale System Design

References

1. History of Fingerprinting. http://onin.com/fp/fphistory.html
2. Federal Bureau of Investigation: The Science of Fingerprints (Classification and Uses), 12–84 edn. US Government Printing Office, Washington DC (1984)
3. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003)
4. Ashbaugh, D.R.: Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology. CRC Press, West Palm Beach, FL (1999)
5. Chen Yi, Demirkus, M., Jain, A.K.: Pores and ridges: high-resolution fingerprint matching using level 3 features. Trans. Pattern Anal. Mach. Intell. **29**(1), 15–27 (2007)
6. Wilson, C., Garris, M., Watson, C.: Matching performance for the US-VISIT IDENT system using flat fingerprints, NISTIR 7110, May 2004, ftp://sequoyah.nist.gov/pub/nist_internal_reports/ir_7110.pdf
7. Federal Bureau of Investigation: CJIS electronic biometric transmission specification, appendix J, descriptions and field edit specifications For type-9 logical records. www.fbibiospecs.org
8. ANSI/NST-ITL 1–2007: Data Format for the Interchange of Fingerprint, Facial, & Other Biometric Information. http://fingerprint.nist.gov/standard/

Unnecessary Data Collection

A service provider typically requires data from an end-user in order to provide the appropriate type or level of service to the end-user. Data that are not relevant for this purpose are deemed unnecessary, and should be labeled as optional. In principle, the service provider must be able to explain why each and every data item collected is necessary, while the end-user has the right to such an explanation.

► Privacy Issues

Unsupervised

Unsupervised is the class information of data which are not available. Algorithms are designed purely based on the attributes of data and data distribution.

► Fusion, Rank-Level
► Linear Dimension Reduction

Unsupervised Rank Level Fusion

The rank level fusion methods can be generally categorized under the headings of supervised and unsupervised. The rank level fusion method that does not require any training data to achieve the fusion of ranks can be categorized as unsupervised method. Thus, using unsupervised rank level fusion methods, one can combine the ranks from different matchers without any "teacher" or training data. The Highest rank method and Borda count methods are the examples of unsupervised rank level fusion methods.

► Fusion, Rank-Level

Unvoiced Sounds

The unvoiced speech is generated by constriction of the vocal tract narrow enough to cause a turbulent airflow, which results in noise, e.g., in fricatives like /f/, /s/, or breathy voice (where the constriction is in the glottis). Unvoiced plosives like /p/, /t/, /k/ fall into this category, too.

► Speech Production

Usability

The extent to which a product can be used by specified users to achieve specified goals. Usability testing employs techniques to collect empirical data during the observation of users using the product for a specific task in order to rectify usability deficiencies of a product. The ISO document 924111 discusses three factors that compose usability: effectiveness, efficiency, and satisfaction. The IEEE Standard Computer Dictionary further describes usability as the "ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component."

► Ergonomic Design for Biometric Systems
► Hand Geometry

User Acceptance

Marek Rejman-Greene
Home Office Scientific Development Branch, Sandridge, St Albans, Herts, UK

Definition

User acceptance: the demonstrated willingness within a user group to employ information technology for the tasks it is designed to support [1].

Introduction

The effective use of many of the applications that include a biometric component requires users to follow specified procedures, and hence, calls for cooperation from the user. Such cooperation is predicated on the acceptance of the biometric technology. In recent years, there have been a number of studies addressing the nature and extent of such acceptance, although our understanding is still partial and further research is needed. Among the often cited reasons for willingness to use a biometric system is trust in the technology as well as in the organization holding biometric reference data – although many of the major studies have not explored user acceptance in sufficient depth to elicit the reasons for a lack of acceptance among a minority of the population.

In a wider sense, user acceptance of IT technologies has been the subject of research for more than 20 years, as ever more complex systems were developed for use in personal, corporate, and government applications [2]. Unfortunately, many of the documented case studies are limited to self-reporting by survey respondents rather than the observation of usage (e.g., obtaining take-up metrics), and often research is addressed to systems where the use of a technology is voluntary. The validation of models such as TAM, the Technology Acceptance Model of Fred Davis [3] (and its successors), remains to be demonstrated as each new technology tests their assumptions.

In his original paper, Davis identified two determinants of an individual's intent to use an IT system: perceived usefulness and perceived ease of use. Subsequent developments of the model have also included aspects such as subjective norm (the subject is influenced by the positive attitudes of people in

positions of respect), personal experience, and demonstrability of the success of the new technology in achieving a goal, as well as the extent to which the use of the application enhances the subject's status or image [2].

Even though there is considerable empirical support for this type of modeling in the deployment of other IT components, such modeling has yet to be validated for user acceptance of biometrically enabled applications. Nevertheless, the literature on user acceptance of biometric systems offers numerous indicators of intent to use – or not use, prominent among which are concerns about privacy and data misuse, health risks, usability issues, and uncertainty about system reliability.

For some applications, especially where convenience is the principal reason for using biometrics, user acceptance may not be sufficient. In these instances, the challenge is to recast the metrics for success in terms of the users' *satisfaction* with the application (and, hence, with the biometric elements). Furthermore, we can envisage systems where the use of the biometric is so integrated with an application that it is no longer seen to be exceptional, for example, in multiplayer, multimedia games where verification by facial image and/or speech blends into the player's immersive environment.

Historical Context

Biometric systems have been deployed for over 30 years, yet the literature to support a unified view on their acceptance has been sparse. Early trials of the technology added a short questionnaire for participants in scenario tests, asking about acceptance and comfort in use, without a more detailed exploration of the reasons for any concerns.

A 1995 review of user acceptance studies summarizes the position at that time [4]. This notes the groundbreaking 1988 study for the State of California's Department of Motor Vehicles on retinal scanning and fingerprint technologies. In this extensive study, considerably more participants declined to use the eye method than the live scan fingerprint equipment – perhaps related to rumors circulating at the time about the risk of disease transmission through eye fluids.

Reference is also made in this review to the results of the 1991 Sandia National Laboratories' pioneering performance evaluation of five biometric modalities [5]. Nearly 100 employees and contractors took part in the trial in "an office-like environment." Of these, 76 returned a questionnaire that sought their views on matters that have appeared time and again in subsequent studies: which devices required most concentration and most proficiency? which were most frustrating to use? and which gave health and safety or privacy concerns or were most intimidating? The hand geometry system was rated as the most acceptable in most of the categories, while a retinal scanning system and one of the two voice systems were viewed as the least acceptable.

Biometric technologies can be applied in a wide range of systems: from personal use in mobile handsets and laptops, and access control at doors of homes and cars, through to the management of time and attendance at work and the control of migration at national borders. With such a diversity of uses, user acceptance is likely to vary considerably according to the proposed application and its context, since concerns that have been raised in studies, such as "invasion of privacy," trust in the quality of the database management, perceptions of health risks [6] etc., are likely to differ in prominence for each context.

More recently, as large-scale applications are deployed, the trend has been to use telephone surveys, even though respondents are unlikely to have experienced the range of biometric modalities about which they are questioned, nor they had the opportunity to view this type of technology [7, 8]. Although these surveys (often with over 1,000 participants) can provide valuable insights into the *perception* of biometric technologies across wider populations, focus group studies will be needed to elicit the rationale behind the statistics and comment on specific issues based upon users' experience of the equipment (and its application) at first hand [6, 9].

Acceptance, as reported in surveys, will depend on the context in which the question is put and the trust that the respondents place in the integrity of the survey organizers. As a consequence, reports that claim that a *specific biometric technology* is more or less acceptable to a population need to be examined critically. There should always be supporting evidence of a validated methodology in the context of one or more types of

application. As a minimum, a study should offer the following in order to interpret its results:

1. A sample questionnaire and the associated telephone scripts, especially if the questions relating to biometrics form part of a more general survey (scripts may describe certain features of the technology, while remaining silent about crucial aspects of their operation);
2. The dates of beginning and completion of the survey, as media comments during the survey period may affect the results;
3. The method of selection of participants, the numbers successfully contacted, together with an indication of reasons for nonresponse and any weighting factors applied to the results; and
4. Any incentives offered

It is expensive to undertake combined scenario trials and facilitated focus groups, along with extensive questionnaire surveys, even though this strategy is likely to give the best indication of the user acceptance of systems to be deployed in the future [9]. However, their relevance begins to diminish with time as better biometric equipment and user interfaces are developed, and the media debate moves on to other aspects of the deployment.

The Importance of User Acceptance

Although there has been limited research to prove the point, it is an established axiom in the field that confident and cooperative users of biometric technology are a prerequisite for successful applications. Sasse [10] notes the close link with usability of systems, since systems that are difficult to use, or where users have unresolved concerns, will result in more verification errors, longer throughput times, and therefore, lead to additional costs for the operator of the system. Reduced usability has also been associated with a reduction in the trust users place in a service.

Reference is also made by Sasse to the Biovision Technology Roadmap [11], which concluded that three factors lead to the acceptance of biometric technology: trust in the security enhancement offered by this form of authentication; greater convenience in use compared with alternative systems; and trust in those holding the biometric data, maintaining the security of that data and not using data for other purposes. In addition to the absence of these factors, other reasons that lead to systems being less accepted include fears of health effects of long-term use of the equipment and the reliability of the identification process.

Principal Large Scale Studies

In 2001–2002, ORC International undertook two nationwide telephone surveys (commissioned by the National Consortium for Justice Information and Statistics) to assess attitudes toward the use of biometrics by government and the private sector [7]. As the first was conducted directly after the events of 9/11, a follow up survey was needed a year later to ensure that responses were not colored by the extensive media discussion of the response to terrorist attacks. The authors of the studies commented, however, that very little had changed in the intervening time. Over 1,000 US adults were sampled in each year and the full results of the survey, together with sample questionnaires and telephone scripts, are available in the public domain.

The study found that only half of the sampled population was aware of the existence of biometrics, indicating that public acceptance would require further education and marketing initiatives. Personal experience of its use was also very low: only 5% of the sample (57 users) had ever used a biometric system, with the majority reported as feeling comfortable with the experience.

Support for the use of biometrics in major government systems (passport verification, entry to government buildings, and for airport check-in) was above 80% in 2002, although these figures represented a decrease in acceptance from the immediate post 9/11 period. For applications designed to counter terrorism, about two-thirds of the survey participants trusted that there would be no unwarranted extension of their use ("function creep"). Nevertheless, there was strong support for privacy safeguards in accordance with Fair Information Practices. Other government applications that scored highly were those in support of law enforcement and reduction of social security fraud.

There was also clear support for private sector use of biometrics, although expressed with markedly less enthusiasm. Respondents voiced support for biometric

verification in the selling of guns and for credit card transactions that involved large sums of money.

In March 2005, TNS and TRUSTe commissioned an Internet survey of attitudes to the application of biometrics [12]. Nearly three years after the ORC study, the awareness of biometrics in the US population had risen to 75%. As in the 2001–2002 study, substantial majorities supported the government use of these technologies, although a group of 15–20% of subjects were opposed to some of the more contentious applications such as use by employers for identity verification and in national identity systems. A comparison of US with Canadian respondents in this research showed that use in the private sector is viewed with more suspicion in Canada (e.g., 55% in the US would support the use of biometrics in employer identity schemes in contrast to 36% of Canadians).

In a question about the acceptability of different modalities for proving identity, fingerprint recognition was a clear leader (81%, with 58% for iris methods appearing in second place).

US respondents were also asked about some of the negative aspects of deploying biometrics at scale. The technologies themselves were trusted, but they were seen as expensive and likely to be defeated by criminals; a significant proportion would not trust governments to limit the use of biometric data to originally stated aims.

In 2004, UKPS, the UK Passport Service, analyzed attitudes to the use of biometrics as part of a very large scale enrolment trial for those biometric technologies that had been proposed for the National Identity Scheme: facial, fingerprint, and iris recognition [13]. With a total of more than 10,000 participants, this is one of the largest trials ever undertaken. Two thousand individuals formed a quota sample, representing the diversity of the UK population, with an additional 750 subjects with a range of disabilities. The remaining users included some who applied directly to participate in the trial. In addition to testing the processes of biometric enrolment, a major goal was to assess the customer experience of recording biometric features.

The overwhelming majority of participants in the UKPS trial found that, overall, the experience was at least as good as they had anticipated, although the iris recognition system experience was noted as least satisfactory, comments being made about the need to remain still and wait for a long time. Advances in technology since 2004 have aimed to address both of these criticisms.

In general, the majority of the group of trialists was not overly concerned about the recording of biometrics (with the exception of comments by disabled people about the use of iris systems). Concerns were lessened after users enrolled into the system. It was notable, however, that greatest concern regarding the use of biometrics was expressed by the Black and Minority Ethnic group and by subjects aged between 18 and 34.

In the quota group, iris recognition was the preferred biometric for males (iris as a first choice for 51%, with a second preference for fingerprints at 24%), female participants preferring this modality to fingerprint use (45% against 36%).

In answer to the question of whether they were in favor of biometrics being used as a means of identification for passport purposes, over 90% of the quota group were either "in favor" or "strongly in favor." In line with their concerns about iris biometrics, disabled people were somewhat less in favor of its use in this context.

In September 2008, the Lieberman Research Group conducted a survey in a number of European countries to ascertain concerns about national and personal financial security, security of dealing on the Internet, and people's personal safety [8]. For the first time, this annual survey included a supplementary question about biometrics, enabling a comparison of attitudes toward the use of these technologies across a number of European countries. The question was posed as: "Which of the following (. . . *methods of authentication*. . .) would you be willing to use to verify your identity with banks and government and other organizations to prevent fraudulent misuse of your personal information?" The interpretation of the results has to be viewed against the generally threatening world, of which participants were reminded in the earlier part of the survey, and the lack of any further information regarding the technologies themselves – or the participants' familiarity with them. The question also brings together use by both public and private sectors, whereas earlier studies have shown significant differences in acceptance.

In all the seven European countries surveyed, the order of acceptance of biometric technologies is fingerprint, eye scan, and voice. "A scan of blood vessels in your hand" was least acceptable of the five modalities. Some countries appear to be noticeably more willing to

use biometrics. For example, willingness to use fingerprint systems is highest in Netherlands (80%) and the UK (75%), and lowest in France (58%). The surveys picked up some significant differences in user acceptance for certain age groups and between males and females in a number of countries. For example, in the Netherlands, less than 50% of the 18–34 age group were willing to consider offering a face image for verification, whereas two-thirds of the over 65s would, in principle, find this acceptable. However, caution should be exercised in interpreting such trends, since the survey respondents' mental models of biometric technologies could be at variance with reality.

Technologies

Smaller scale studies have attempted to understand the issues of user acceptance in respect of individual biometric technologies. In several of the major studies, fingerprint technology has been rated as most acceptable. For example, in the Trustguide focus group studies [6], it was seen as the "least invasive and the most acceptable form of biometric identification." The report authors mention that this could be context-specific, and may reflect the perception of high accuracy in its use in the detection of crime.

In a study of the use of fingerprint biometrics, researchers from the Canadian National Research Council questioned a sample of 24 individuals in order to understand some of the considerations that might be influencing comments by users [14]. These subjects were firstly asked to simulate a number of Internet purchases, using biometrically secured personal or corporate credit cards. In the discussion afterwards, there was considerable confusion about the security value of biometrics, but users felt that they were likely to accept the technology once it became more prevalent. Privacy concerns were mentioned without being rated as overly important. However, as the facilitator probed more deeply, subjects began to be less certain of their acceptance. Even though this was a limited study, it points to the need for more detailed investigation of users' perceptions and willingness to use these modalities.

Occasionally, comments have been made in the literature about the fears of disease transmission, as users touch the surface of conventional fingerprint sensors one after another. In some deployments, cleaning tissues are provided, while users are reminded of the numerous occasions on which they touch common surfaces such as door handles and terminals. For a small number of deployments, these assurances may not be sufficient. Munyan at Computer Sciences Corporation reports that the immigration control authorities in one Asian country required the decontamination of the fingerprint sensor before every use by travelers [15].

In the evaluations of user acceptance described earlier, voice verification has not been rated highly in spite of its growing application in corporate password reset systems. Attitudes to its possible use were explored in a Harris Interactive telephone poll conducted in April-May 2008 with 553 UK residents who had been in contact with a Customer Service Center during the previous year [16]. A description of the process of enrollment resulted in only 38% of subjects rating it as a technology they would be likely to use, even though 60% were confident that it was secure. By listening to a demonstration, the willingness to accept voice biometrics increased to 51%. (As in services using other biometric modalities, users seem to need reassurance about the impact of verification failure on their everyday lives.) Additional data in this survey suggest that those who did not accept speaker recognition were also concerned about unrelated aspects of automation of the phone interaction with the service provider.

Strategies for managing the handover between a failed speaker verification session and a human operator have been examined in an earlier UK study [17]. Two-hundred and seventy onesubjects were authenticated over the public telephone network. From the options investigated, subjects preferred the protocol of a message: "I'm sorry the voice verification process has not been successful, please hold while I connect you to a human operator" after *two* failed machine verifications. Research with a small number of subjects in the US hints at the need for operators to communicate trust during any speaker verification dialog [18].

To assess the likely acceptance of a biometric technology in the context of an application, focus groups can be used at first to elicit major concerns. This qualitative research should inform the design of questionnaires that can then track changes in acceptance, as individuals experience the application in both in-house and public-facing trials.

A successful six-month trial of iris recognition in a banking (public ATM) context demonstrates this approach [9]. At the beginning, focus groups identified

issues of trust, questioning whether such a "futuristic" technology could be used in consumer financial transactions, as well as noting health concerns. Participants in the groups were also invited to review proposals for a marketing campaign. In the second (questionnaire) stage, concerns with reliability, health aspects, and data misuse surfaced – even before participants were exposed to the technology. As users became familiar with the use of iris biometrics, progressively higher levels of acceptance were recorded as users enrolled in the system, used an early prototype to withdraw cash, and then experienced the final prototype.

Subsequently, the 1998 public trial in Swindon, UK (with about 400 users from the general public) confirmed this progression: from 44% comfort levels directly after iris enrolment to 94% after use of the IRIS ATM.

Applications

Biometric technologies are integrated into applications that deliver a service to an operator or individual, and users may not separate the biometric aspects from concerns about the application [19]. In such cases, the results of user acceptance testing will be partly determined by the biometric technology and partly by the application. For example, in a 2008 pilot for the UK's Identity Card for Third Country Nationals [20], a high degree of user satisfaction was recorded for the process as a whole. Over 12,000 sets of facial images and ten fingerprints were recorded, stored, and matched. The analysis of feedback forms showed that in excess of 90% of the customers rated the enrollment service as satisfactory. Dissatisfaction was registered by 8% of the group – a figure that challenges the service designers to improve the complete customer experience.

Summary

For many applications with a biometric element, user acceptance is likely to be a key issue for their successful operation. However, there has been an absence of systematic modeling of the nature of user acceptance in such systems. Focus groups and other qualitative research have identified some of the key determinants; trust in the technology, fears of misuse of biometric data, and health concerns have been mentioned.

It is unfortunate that the few larger-scale telephone surveys on user acceptance have not followed up these observations with more targeted questions. In any case, few of the participants in these surveys would have had any first hand experience of using a biometric system – a key factor in improving user comfort and acceptance of these novel technologies as evidenced in the UKPS Biometric Enrollment and IRIS ATM trials.

Further research should be associated with the design, development, and piloting of new biometrically enabled applications, with the aim of identifying the main factors influencing user acceptance of technology and application, both at enrollment and in service use. Since the usability of a biometric system is believed to play a substantial role in user acceptance, research targeting both the aspects is recommended. However, in the longer term, the goal should not be mere acceptance, but a positive degree of satisfaction with the end-to-end service.

Related Entries

- ▶ Privacy
- ▶ Testing
- ▶ Usability

References

1. Dillon, A.: User acceptance of information technology: theories and models. In: Williams, M. (ed.) Annual Review of Information Science and Technology, vol. 31, pp. 3–32 (1996)
2. Venkatesh, V., et al.: User acceptance of information technology: toward a unified view. MIS Quart. **27**, 425–478 (2003)
3. Davis, F.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart. **13**, 319–339 (1989)
4. Survey: so what do people think of biometrics? Biometrics Technology. Today **3**, 7–10 (1995)
5. Holmes, J.P., et al.: A performance evaluation of biometric identification devices. http://infoserve.sandia.gov/cgi-bin/techlib/access-control.pl/1991/910276.pdf (1991)
6. Lacohee, H., et al.: Trustguide: final report. www.trustguide.org.uk (2006). Last Accessed 24 March, 2009
7. ORC International: attitudes toward the uses of biometric identification technologies by government and the private sector. http://www.search.org/files/pdf/Biometricsurveyfindings.pdf (2002). Last Accessed 24 March, 2009
8. Unisys Security Index – November 2008 (Wave III). http://www.unisyssecurityindex.com/resources/viewpoints/Executive%20summaryfinal.doc (2008). Last Accessed 24 March, 2009
9. Coventry, L., et al.: Usability and biometric verification at the ATM interface. CHI Lett. **5**(1), 153–160 http://research.iiit.ac.in/

~vandana/PAPERS/BASIC/large-scale-ATM.pdf (2003). Last Accessed 24 March, 2009
10. Biovision consortium: Biovision: roadmap for biometrics In Europe to 2010. http://ftp.cwi.nl/CWIreports/PNA/PNA-E0303.pdf (2003). Last Accessed 24 March, 2009
11. TNS/TRUSTe: Consumer attitudes about biometrics in ID documents. http://www.truste.org/pdf/Biometrics_Study.pdf (2005). Last Accessed 24 March, 2009
12. Atos Origin: UK Passport Service Biometrics Enrolment Trial. http://www.ips.gov.uk/passport/downloads/UKPSBiometrics_Enrolment_Trial_Report-Management_Summary.pdf (2005). Last Accessed 24 March, 2009
13. Sasse, M.A.: Usability and user acceptance of biometrics. http://www.cesg.gov.uk/policy_technologies/biometrics/media/usability_and_user_acceptance.pdf
14. Heckle, R.R., et al.: Perception and acceptance of fingerprint biometric technology. In: Proceedings of the Symposium on Usable Privacy and Security (SOUPS) National Research Council Canada. http://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-49363.pdf (2007). Last Accessed 24 March, 2009
15. Munyan, D.: Coming clean on hygiene. Biometrics Technology. Today **13**(4), 7–8 (2005)
16. Nuance: Understanding consumer perspectives on voice biometrics. Corporate publication (2008)
17. Dialogues Spotlight Research Team: dialogues for speaker verification/operator handover. http://spotlight.ccir.ed.ac.uk/public_documents/technology_reports/No.5%20Verification.pdf (2000). Last Accessed 24 March, 2009
18. Turner, C.W., et al.: The effects of use on acceptance and trust in voice authentication technology. In: Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, Human Factors & Ergonomics Society, Santa, Monica, CA, pp. 718–722 http://triangleinnovation.com/Documents/HFES2006VoiceAuthTrust.pdf (2006)
19. Patrick A.: Acceptance of biometrics: things that matter that we are ignoring. In: International Workshop on Usability and Biometrics, NIST: Washington, D.C. http://zing.ncsl.nist.gov/biousa/docs/workshop08/day1/7Patrick/Andrew-Patrick-Acceptance-of-Biometrics.pdf (2008). Last Accessed 24 March, 2009
20. UK Border Agency: Biometric enrolment pilot evaluation – executive summary. http://www.ukba.homeoffice.gov.uk/sitecontent/documents/managingourborders/biometricenrolmentevaluation (2008). Last Accessed 24 March, 2009

User Interfaces

User interfaces are systems (including hardware and software components) that facilitate the interaction of users with the system.

► Biometric Systems, Agent-Based

User Interface, System Design

Jianjie Li, Xin Yang, Xunqiang Tao, Jie Tian
Institute of Automation, Chinese Academy of Sciences, Beijing, People's Republic of China

Synonyms

Interface; Graphical user interface

Definition

The User Interfaces have been around as long as computers have existed, even well before the field of Human–Computer Interaction was established [1, 2]. The user interface provides means of: input, allowing the users to manipulate a system; output, allowing the system to produce the effects of the users' manipulation [3].

Introduction

Computer software has become pervasive in today's society. How to design and create easy usable software is the central issue in software development. As new software products are developed the emphasis has often been on what features and function they contain rather than how the features are used. This emphasis is often reflected as user interface. The user interfaces are the main and dispensable components of any software [2].

The user interface is often used in the context of computer systems and electronic devices. The user interface of a mechanical system, a vehicle, or an industrial installation is sometimes referred to as the Human–Machine Interface (HMI) [4]. HMI is a modification of the original term Man–Machine Interface (MMI). In practice, the abbreviation MMI is still frequently used, although some may claim that MMI stands for something different now. Another abbreviation is Human–computer interaction (HCI) [5], but is more commonly used for Human–computer interaction than Human–computer interface. In a word, the terms refer to the "layer" that separates a human who is operating a machine from the machine itself [3].

When designing human/computer systems, the user interface between human and system is crucial. It is the communication channel that end-user can interact with a system [5]. A good user interface is to find out user tasks as one of the first steps and driven in large part by human aesthetics in their look and feel [6]. Users have to be able to control the system and access the state of the system. For example, when driving an automobile, the driver uses the steering wheel to control the direction of the vehicle, and the accelerator pedal, brake pedal, and gear stick to control the speed of the vehicle. The driver perceives the position of the vehicle by looking through the windscreen and exact speed of the vehicle by reading the speedometer. The user interface of the automobile is the instruments that the driver can use to accomplish the tasks of driving and maintaining the automobile [2]. In a biometric' system design, the user interface refers the aggregate designed into an information device with which a human being may interact. The user interface should provide means of "Input" which allows the users to manipulate a system and "Output" which allows the system to produce the effects of the users' manipulation.

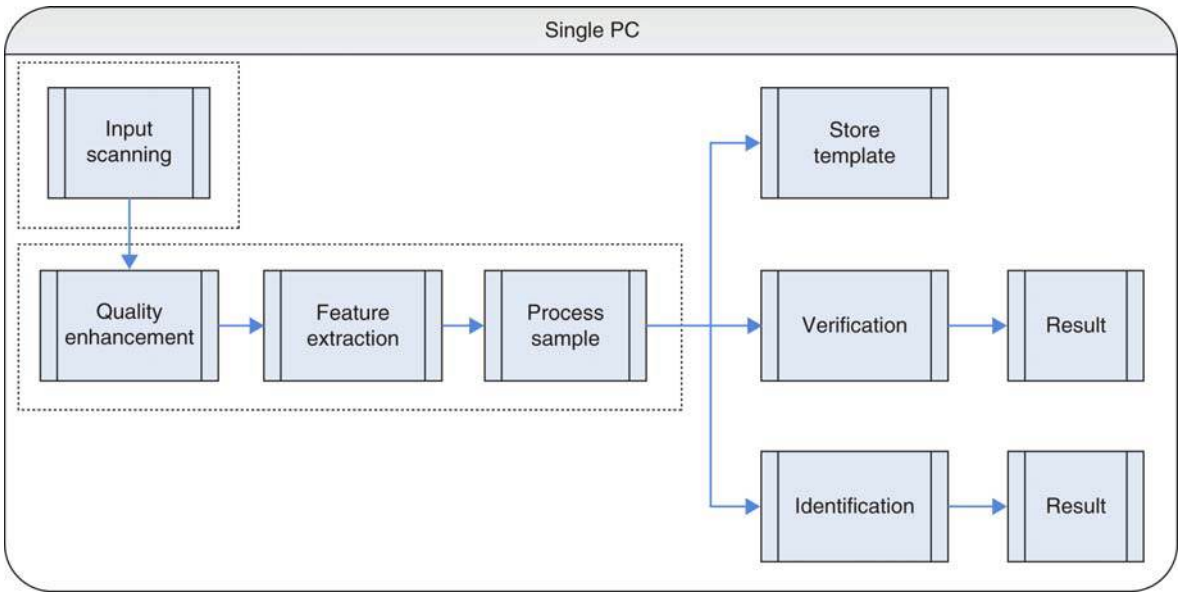
User interface design is an expensive, complex, and time consuming process [7]. To provide an environment where developers can design and implement user interfaces in a professional and systematic way, there are a lot of good software to help developers design and implement the user interface such as QT and MFC.

In biometric system design, the user interface is a little different from other system design. Biometric system always has capture device for input biometric feature. The graphical user interface should have a picture control to show the image of the capture, and the user can discard the image because of poor quality. The biometric system need to set the parameter for different purpose. So user interface need to have some interaction with the user. There is two kind of system, one for Single mode and the other for Client/Server mode.

Single Mode

The single mode means all the system run on a single PC. That is a common mode. It is simple in system design, easy to install, and undertaken on a small scale. Figure 1 shows the architecture of the single mode. When system works in single mode, UI can be described as follows:

1. User interface include a page or dialog including some exit boxes or scroll bars that can set the configuration of the system. This step initializes the biometric system.



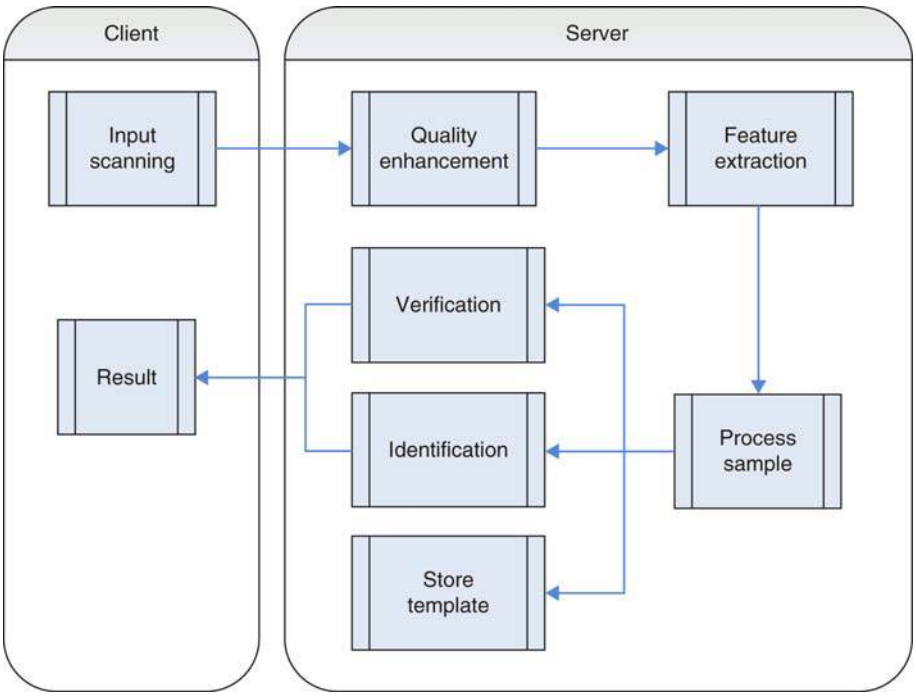
User Interface, System Design. Figure 1 Architecture of the Single mode.

2. User interface include the button like "connect the capture" to make ready of the capture, then call the capture's API to try to open the sensor or camera. If the capture is ready, the program opens the corresponding capture interface. Then a biometric feature can be captured. If any error happens, the program return a message box or show a message on the dialog, which let the user know that the capture is not ready. The user interface also shows where the error occurred, which help the user to find the problem.
3. When the biometric feature is captured, the user interface shows the image to the user. Note that a good automated image mechanism ought to be employed so that the system will be able to obtain an image of sufficient quality. If the image does not have enough quality, the user interface should cancel current image, and let the user capture the biometric feature again.
4. There is a push-button interface to select from two applications: Enrollment and Match. In enrollment process, user should input other information like ID number, age, sex, etc. That may need several edit boxes, scroll bars, and combo boxes. After that, a message should be given to tell the user whether Enrollment or Match succeeds or not.
5. In single mode, the biometric feature does not always have a database. Instead, program store the feature in a folder with some files. And some use the desktop database like Microsoft Access. What ever, the user interface could check the database conveniently if possible.

Client/Server Mode

In client/server mode, there is a high performance computer called server, the algorithms are run in server, and the entire database are stored in server. Figure 2 shows the architecture of the client/server mode. There are server advantages in client/server mode [8].

1. The client PC may not have sufficient power to run the algorithms on real-time.
2. The user database maybe on a server and that is a safe mode.
3. The matching algorithms will execute in a more secure environment.



User Interface, System Design. Figure 2 Architecture of the Client/Server mode.

4. Identification over large populations can only reasonably be done on a server.

The client part is a common PC. It is used for input or scans the biometric feature, show the result from the server.

1. The user interface includes the button to make ready of the capture, then call the capture's API to try to open the sensor or camera. If the capture is ready, the program opens the corresponding capture interface. If some error occurred, the program should return a message box or show a message on the dialog, which let the users know the capture is not ready. The user interface should show where the error occurred, which make the user easy to find the problem.
2. When the biometric feature is captured, the user interface shows the image on the dialog to the user. That can let user make sure the image is suitable for the situation. In some situation, user should also input some information like ID number, age, sexual, etc. If the image does not have enough quality, the user interface should cancel current image, and let the user capture the biometric feature again.
3. The user interface should have a button to connect the server. If the server is ready, then the client could transfer the biometric feature and other information to the server using net protocol, and also tell the server what to do with these in formations. For example, verify with the ID number, or identify in a database. If there are some errors occur in this process, the program should show a message on the dialog in order to let the user know what and where the error is.
4. After the server calculates for a moment, the client will receive a message about the result from the server. The user interface should change the numeric result to an intuitionist message, and show it on the dialog, for example, a message box or an obvious picture. And user could decide whether to capture it again.

The server is a high performance computer. It can store the database; calculate the huge operation in very short time. It has a super user who has the greatest power.

1. The user interface has a log in dialog for user to input the name and password. It also can capture the biometric feature to log in.
2. User interface need to manage the common users. That interface should include a list of all the users. When click one of them, the interface should show all the detail about this user. There is a button to show a dialog, and allow super user change the common user's competence or details. In this dialog, super user can insert or delete a common user.
3. User interface need to manage the parameter of the system and the algorithms. This interface should correspond to the project. It can change the match score parameter to decide if it necessary to reduce the false accept rate (FAR).
4. User interface needs to manage the database. The program connects to the database and has a dialog to show the detail. The super user can delete or insert a record by push the corresponding button.

Summary

Technology alone may not win user acceptance and subsequent marketability. The user experience, or how the user experiences the end product, is the key to acceptance [9]. The importance of good user interface design can be the difference between product acceptance and rejection. So, in biometric' system design good User Interface Design must make a product easy to understand and use, which can result in greater user acceptance.

Related Entries

► Biometric System Design, Overview

References

1. Jorgensen, A.H., Myers, B.A.: User interface history. In: Proceedings of Conference on Human Factors in Computing Systems, Florence, Italy, pp. 2415–2418 (2008)
2. Jean, A., Pratt, Hauser, K., Ugray, Z., Patterson, O.: Looking at Human–Computer Interface Design: Effects of Ethnicity in Computer Agents. Interact. Comput. **19**(4), 512–523 (2007)
3. http://en.wikipedia.org/wiki/User_interface
4. Krepki, R., Curio, G., Blankertz, B., Berlin, K.R.M.: Brain-Computer Interface-The HCI Communication Channel for Discovery. Int. J. Hum. Comput. Stud. **65**(5), 460–477 (2007)

5. Apple. Apple human interface design guidelines: Introduction to apple human interface guidelines (2006)
6. Kamal, A., Asad, I., Bashar, B.A., Alaa.: Teaching human computer interaction methods in embedded system design 2008. In: Proceedings of Third International Conference on Information and Communication Technologies: From Theory to Applications, pp. 1–6. Damascus, Syria (2008)
7. Russo, G., Birtolo, C., Troiano, L.: Generative UI design in SAPI project. In: Proceedings of Conference on Human Factors in Computing Systems, Florence, Italy, pp. 3627–3632 (2008)
8. http://www.usernomics.com/user-interface-design.html
9. BioAPI 2.0, BioAPI Consortium: (2005). http://www.bioapi.org/Version_2.0_Description.asp. Accessed 2005

User-Centered Design

Pheasant summarized ergonomic design by the principle of user-centered design, which "If an object, a system, or environment is intended for human use, then its design should be based upon the physical and mental characteristics of its human users". Moreover Woodson states the design should allow users to complete the desired functions and tasks with minimal stress and maximum efficiency. Therefore, the object of ergonomics and user-centered design is to achieve the best possible match between the product and users in the context of the task to be performed. Chignell and Hancock referred to this as the "design triad," which consists of three primary relationships. The first is the user-task relationship, which is much like task analysis and answers the following questions: What is the task, and how is it carried out by the user? The second relationship is user-artifact, which is the relationship between the user and the system and lies at the heart of ergonomics. Lastly, the artifact-task relationship represents the methodology for using the system to perform the task, which is also known as methods improvement. Other techniques, methods, and practices of User-centered Design besides usability testing and audits include: interviews, focus group research, surveys, design, cognitive, or structured walk-throughs, paper and pencil evaluations, expert evaluations, field studies, and follow-up studies.

▶ Ergonomic Design for Biometric Systems

User-dependent Fusion

▶ Fusion, User-Specific

Utility

Utility is the observed performance of a sample in a biometric system, or similarly the impact of an individual biometric sample on the overall performance of a biometric system. The characteristic of the sample source and the fidelity of the processed samples contribute to or similarly detract from the utility of the sample.

▶ Speech Production

Utterance

Utterance is a spoken input speech sample. It may be real time streaming audio, a prerecorded file, or the result of buffering. In interactive systems, a single utterance generally corresponds to a single interaction turn.

▶ Speaker Recognition, Standardization

V

Vascular Biometrics

► Vascular Image Data Format, Standardization

Vascular Image Data Format, Standardization

Alex H. Choi[1], Jonathan R. Agre[2]
[1]Department of Information Engineering Myongji University, Seoul, South Korea
[2]Fujitsu Laboratories of America College Park, MD, USA

Synonyms

Vascular biometrics; Vein biometrics

Definition

A Vascular ► Biometrics, Overview Image Format Standard is useful for the exchange of vascular biometric image information across different systems developed by multiple organizations. As one part of this standardization effort, the International Standard Organization (ISO) has published a standard for a vascular biometric image interchange format, which is the ISO/IEC 19794-9 (Biometric Data Interchange Format – Part 9 Vascular Image Format). The standard includes general requirements for image capture devices, environmental conditions, specific definitions of image attributes, and the data record format for storing and transmitting vascular biometric images. The vascular biometric image format standard was developed in response to the need for system interoperability which allows different vascular biometric systems to be easily integrated with other biometric modalities in a large-scale system.

Introduction

Vascular biometric technologies have existed for many years. Moreover, new technologies employing vascular images obtained from various parts of the human body are emerging or under continuous improvement as a result of new, state-of-the-art imaging devices. Some of these technologies are being widely adopted as reliable biometric modalities [1].

Vascular biometrics offer several intrinsic advantages in comparison with the other popular biometric modalities. First, the vascular imaging devices use near-infrared or infrared images to capture the vein pattern underneath the skin. This provides a high degree of privacy that is not available with fingerprints, which can be unintentionally left on objects, or by facial images for face recognition schemes, which are easily captured without ones knowledge. A similar possibility exists for iris images captured without consent for use in iris recognition schemes. Second, the vascular imaging devices can be constructed to operate in a non-contact fashion so that, it is not necessary for the individual to touch the sensor in order to provide the biometric data. This is advantageous in applications that require a high degree of hygiene such as medical operating room access or where persons are sensitive about touching a biometric sensing device. Third, a high percentage of the population is able to provide viable vascular images for use in biometric identification, increasing ► usability by providing an additional way to identify persons not able to provide fingerprints or other biometric modal data. Fourth, depending on the particular wavelength of (near-) infrared light that is used, the image can capture only the vein patterns containing oxygen depleted blood. This can be a good indication that the biometric image is from a live person. Fifth, the complexity of the vascular image can be controlled so that the underlying amount of information contained in the image can be quite high when compared to a fingerprint, allowing one to reduce the false accept or false reject rates to low

© 2009 Springer Science+Business Media, LLC

levels. At the same time, the image information can be compressed or it can be processed into a template to reduce storage requirements.

Vascular biometric technologies are being used or proposed for many applications. Some of these include access control to secure areas, employee time-clock tracking, Automatic Teller Machines (ATMs), secure computer login, person identification, and as one of several biometrics in multi-biometric systems. The technology is not appropriate for certain other applications such as criminal forensics or surveillance.

Currently, however, little vascular biometric image information is being exchanged between the equipment and devices from different vendors. This is due in part to the lack of standards relating to interoperability of vascular biometric technology. In the general area of biometrics interoperability, the International Standard Organization (ISO) and the regional organizations, such the INCITS M1 group in the US, define a collection of standards relating to the various biometric modalities that include data interchange formats, conformance testing of image and template interchange formats, performance testing and application profiles. The most critical are the formats for information exchange that would ensure interoperability among the various vendors. Definition and standardization of the data structures for the interoperable use of biometric data among organizations is addressed in the ISO/IEC 19794 series [2], which is the multipart biometric data interchange format standard, which describes standards for capturing, exchanging, and transferring different biometric data from personal characteristics such as voice, or properties of parts of the body like face, iris, fingerprint, hand geometry, or vascular patterns.

To address this short-coming in the vascular domain, the ISO has published a standard for a vascular biometric image interface format, entitled the ISO/IEC 19794-9 (Biometric data interchange format – part 9 Vascular image format) [3].

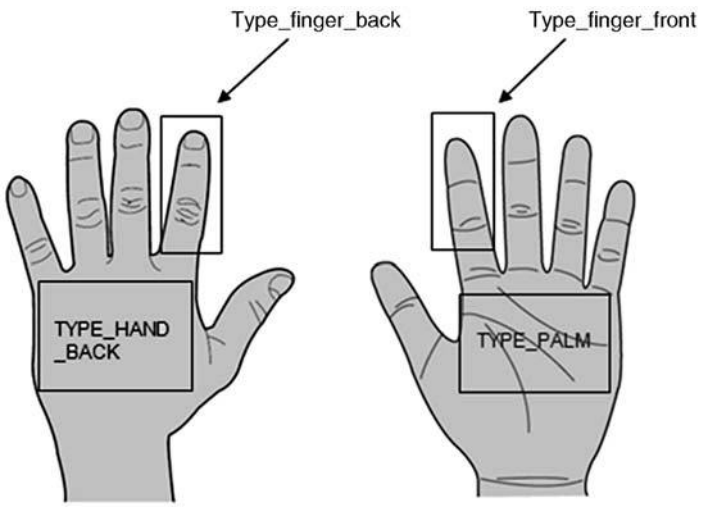
The main purpose of this standard is to define a data record format for storing and transmitting vascular biometric images and certain of their attributes for applications requiring the exchange of raw or processed vascular biometric images. It is intended for applications not severely limited by the amount of storage required and is a compromise or a trade-off between the resources required for data storage or transmission and the potential for improved data quality/accuracy. Basically, it enables various preprocessing or matching algorithms to identify and verify the type of vascular biometric image data transferred from other image sources and to allow operations on the data. The currently available vascular biometric technologies that are commercialized and that may utilize this standard for image exchange are technologies that use the back of the hand, the palm, and the finger [4–6]. There is the ability to extend the standard to accommodate other portions of the body if the appropriate technology is brought forward.

The use of standardized source images can provide interoperability among and between vendors relying on various different recognition or verification algorithms. Moreover, the format standard will offer the developer more freedom in choosing or combining matching algorithm technology. This also helps application developers focus on their application domain without concern about variations in how the vascular biometric data was processed in the vascular biometric modalities.

Introduction to ISO/IEC 19794-9 Vascular Image Data Format Standard

ISO published the ISO/IEC 19794-9 Vascular Image Data Format Standard in 2007, as a part of the ISO/IEC 19794 series. The ISO/IEC 19794-9 vascular image data format standard specifies an image interchange format for biometric person identification or verification technologies that utilize human vascular biometric images and may be used for the exchange and comparison of vascular image data [7]. It specifies a data record format for storing, recording, and transmitting vascular biometric information from one or more areas of the human body. It defines the contents, format, and units of measurement for the image exchange. The format consists of mandatory and optional items, including scanning parameters, compressed or uncompressed image specifications, and vendor-specific information.

The ISO/IEC 19794-9 vascular image data format standard describes the data interchange format for three different vascular biometric technologies utilizing different parts of the hand including back-of-hand, finger, and palm. The standard also supports room for extension to other vascular biometrics on other parts of the human body, if needed. Figure 1 shows an



Vascular Image Data Format, Standardization. **Figure 1** Examples of vascular biometric areas on different parts of the hand [3].

example of vascular biometric areas on different parts of the hand that are specified in ISO/IEC 19794-9.

The interchange format follows the standard data conventions of the 19794 series of standards such as requiring all multi-byte data to be in big-endian format, transmitting the most significant byte first and the least significant byte last, and within a byte, the order of transmission shall be the most significant bit first and the least significant bit last. All numeric values are treated as unsigned integers of fixed-length.

The vascular pattern biometric technologies currently available employ images from the finger, back of the hand, and palm side of the hand. The location used for imaging is to be specified in the format. To further specify the locations, the object (target body) coordinate system for each vascular technology is defined. Standard poses and object coordinate systems are also defined. All the coordinate systems are right-handed Euclidian coordinate systems. It is then possible to optionally specify a rotation of the object from the standard pose. In order to map the object coordinate system to the image coordinate system without further translation, an *x*- and *y*-axis origin for scanning can be specified in the data.

The image is acquired by scanning a rectangular region of interest of a human body from the upper left corner to the lower right in raster scan order, that is, along the *x*-axis from top to bottom in the *y* direction. The vascular image data can be stored either in a raw or compressed format. In a raw format, the image is represented by a rectangular array of ▶ pixels with specified numbers of columns and rows. Images can also be stored using one of the specified lossless or lossy compression methods, resulting in compressed image data. The allowable compression methods include the JPEG [8], JPEG2000 [9], and JPEG LS [10]. It is recommended that the compression ratio be less than a factor of 4:1 in order to maintain a quality level necessary for further processing.

Image capture requirements are dependent on various factors such as the type of application, the available amount of raw pixel information to be retained or exchanged, and the targeted performance. Another factor to consider as a requirement for vascular biometric imaging is that the physical size of the target body area where an application captures an image for the extraction of vascular pattern data may vary substantially (unlike other biometric modalities).

The image capture requirements also define a set of additional attributes for the capture devices such as ▶ gray scale depth, ▶ illumination source, horizontal and vertical resolution (in pixels per cm), and the aspect ratio. For most of the available vascular biometric technologies, the gray scale depth of the image ranges up to 128 gray scale levels, but may, if required, utilize two or more bytes per gray scale value instead of one. The illumination sources used in a typical vascular biometric system are near-infrared wavelengths in the range of approximately 700–1200 nm infrared light sources. However, near-infrared, mid-infrared, and

Vascular Image Data Format, Standardization. Table 1 Vascular image biometric data block

Bytes	Type	Content	Description
1–26		Data block header	Header used by all vascular biometric image providers. Information on format version, capture device ID, number of vascular images contained in the data block, etc.
27–58		Vascular image header	Image header for the first image. Contains all individual image specific information
	Unsigned char	Image data	
		•	•
		•	•
		Vascular image header	Image header for the last image
	Unsigned char	Image data	

visible light sources can be defined and more than one source may be employed.

Table 1 shows the basic structure of the vascular image biometric data block. A single data block starts with a vascular image record header, which contains general information about the data block such as the identification of the image capture device and the format version. One or more vascular image blocks follow the record header. Each image block consists of an image header and raw or compressed image data. The image header contains all the image specific information such as the body location, rotation angle, and imaging conditions. All images in a data block must come from the same capture device. If multiple devices are used, then multiple blocks must be used.

The vascular image record header consist of general information on the vascular images contained in the data block, such as the format version number, total length of the record block, capture device identification, and the number of images contained in the data block. More specific information includes format identifier, version number, record length, capture device ID, and number of images.

For each image in the data block, the vascular image header describes individual image-specific information including image type, vascular image record length, image width and height, gray scale depth, image position, property bit field, and rotation angle. Other information in the vascular image header may include Image format, illumination type, image background, horizontal scan resolution, vertical scan resolution, pixel aspect ratio, and vascular image header constants. The image data follows and is used to store the biometric image information in the specific format defined in the vascular image record header.

Future Activities

There are considerable ongoing standardization activities relating to vascular biometrics, building upon the biometric data interchange format for vascular images standard. A companion document that specifies the conformance testing for the data interchange format is currently under development. The conformance standard specifies how to check whether the data produced by a vascular imaging device, does indeed agree with the interchange format, as well as which items are mandatory or optional. There are also ongoing efforts, both internationally and in the U.S., to include the vascular image formats into the various application profiles (such as the INCITS M1 Profile for Point-of-Sale Biometric Identification/Verification), which define how to use vascular biometrics in the specific context of an application. There are also efforts at including vascular methods in multi-biometric fusion schemes or as a biometric component of a smart-card based solution. Eventually, it is expected that vascular methods will become one of the important biometric modalities, offering benefits not provided by the other techniques in certain applications.

Summary

Vascular biometric technologies including vascular images from the back-of-hand, finger, and palm are being used as a security integrated solution in many applications. The need for ease of exchanging and transferring vascular biometric data from biometric recognition devices and applications or between different biometric modalities requires the definition of a

vascular biometrics data format standard. The development of the vascular biometric data interchange format standard also helps to ensure interoperability among the various vendors. This paves the pathway so that vascular biometric technologies can be adopted as a standard security technology which is easily integrated in various ranges of applications.

Related Entries

- ▶ Back-of-hand Vein
- ▶ Finger Data Interchange Format Standardization
- ▶ Finger Vein
- ▶ Palm Vein
- ▶ Vein and Vascular Recognition

References

1. Choi, A.H., Tran, C.N.: Handbook of Biometrics: Hand Vascular Pattern Recognition Technology. Springer, New York (2007)
2. ISO/IEC 19794-1 Information Technology: Biometric Data Interchange Format – Part 1: Framework/reference model
3. ISO/IEC 19794-9 Information Technology: Biometric Data Interchange Format – Part 9: Vascular image data
4. Im, S.K., Park, H.M., Kim, Y.W., Han, S.C., Kim, S.W., Kang, C.H.: Biometric identification system by extracting hand vein patterns. J. Korean Phy. Soc. **38**(3), 268–272 (2001)
5. Miura, N., Nagasaka, A., Miyatake, T.: Feature Extraction of Finger-Vein Patterns Based on Repeated Line Tracking and Its Application to Personal Identification. Mach. Vis. Appl. **15**, 194–203 (2004)
6. Watanabe, M., Endoh, T., Shiohara, M., Sasaki, S.: Palm vein authentication technology and its applications. In: Proceedings of Biometric Consortium Conference, VA, USA, September 2005
7. Volner, R., Bores, P.: Multi-Biometric techniques, standards activities and experimenting. In: Baltic Electronics Conference, pp. 1–4. Tallinn, Estonia (2006)
8. ISO/IEC 10918 (all parts) Information Technology: Digital Compression and Coding of Continuous Tone Still Images
9. ISO/IEC 15444 (all parts) Information Technology: JPEG 2000 Image Coding System
10. ISO/IEC 14495 (all parts) Information Technology: Lossless and Near-Lossless Compression of Continuous Tone Still Images

Vascular Network Pattern

The network pattern composed of blood vessels.

Human blood vessels develop network structures in each level of artery, arteriole, capillary, venule, and vein. The network of major blood vessels can be seen in funduscopy and in visual observation of body surface. The vascular networks in fundus image are those of retinal arteries and retinal veins. The blood vessels observed on body surface are the cutaneous veins.

Both network patterns can be used in biometric authentication. There are no apparent evidence on the uniqueness and the permanence of the vascular network pattern. However, in practice, the vascular pattern has been used for biometric authentication without a serious problem. Since the retinal pattern is kept inside an eye, it is stable and seldom affected by the change of outer environment. It is not easily observable by others and robust against the theft and the forgery. The retinal pattern is complex, and high identification accuracy can be expected. The authentication using this retinal pattern has been used in the institutions that require high level of security.

The vascular network pattern in a hand and in a finger can be visualized by transillumination imaging or reflection-type imaging using near-infrared light. The authentication with vascular pattern of a hand and a finger is safer and more convenient than that with retinal pattern. It has been used in common security applications such as the authentication in ATM and in access management.

- ▶ Performance Evaluation, Overview

Vascular Recognition

- ▶ Retina Recognition

Vector Quantization

The vector quantization (VQ) is a process of mapping vectors from a large vector space to a finite number of regions in that space (Linde, Y., Buzo, A., Gray, R.: An algorithm for vector quantizer design. IEEE Trans.

Comm. **28**, 84–9517 (1980)). Each region is called a cluster and can be represented by its center called a codeword. The collection of all codewords is called a codebook. During the training phase, a speaker-specific VQ codebook is generated for each known speaker by clustering the corresponding training acoustic vectors. The distance from a vector to the closest codeword of a codebook is called a VQ-distortion. During the recognition phase, an input utterance of an unknown voice is *vector-quantized* using each trained codebook, outputting a VQ distortion for each codebook, each client speaker. The speaker corresponding to the VQ codebook with the smallest distortion is identified. Both for the training and testing phases, the VQ process works independently on each input frame and produces an averaged result (a codebook or VQ distortion). Thus, there is no need to perform a time alignment. The lack of time warping greatly simplifies the system; however, it neglects speaker-dependent temporal information that may be present in prompted phrases.

▶ Speaker Matching

Vein

Veins are the blood vessels that carry blood to the heart. In the cardiovascular system, blood vessels consist of arteries, capillaries, and veins. Veins collect blood from capillaries and carry it toward the heart. In most of the veins, the blood is deoxygenated. The pulmonary vein is one of the exceptions that carry oxygenated blood. The walls of veins are relatively thinner and less elastic than those of arteries. Some veins have one-way flaps called venous valves that prevent blood from flowing back. The valves are found in the veins that carry blood against the force of gravity, especially in the veins of the lower extremities.

The vein in the subcutaneous tissue is called a cutaneous vein. Some of the cutaneous veins can be observed on the body surface with the naked eye. With the light of high transmission through body tissue such as near-infrared light, we can obtain a clear image of the cutaneous vein. Since the pattern of venous network is largely different between individuals, the images can be used for authentication. The biometric authentication using the venous network patterns in a palm and a finger is common.

▶ Palm Vein Image Sensor
▶ Performance Evaluation, Overview

Vein Biometrics

▶ Vascular Image Data Format, Standardization

Vein Recognition

▶ Retina Recognition

Velocity (Speed)

Velocity of pen movement during the signing process.

Velocity features seem to be one of the most useful features of on-line signatures. Generally, velocity is computed from the first-order derivative of the pen position signal with respect to time. The easiest way to compute the velocity is to calculate the distance between two consecutive pen-tip positions if the data is acquired at equidistant sample points. Velocity features are represented in two ways: velocities along the x-axis and y-axis or velocity along the pen movement direction (tangential direction). In the latter case, the direction of pen movement is also considered as a separate feature.

▶ Signature Recognition

Verification

Biometric verification is a process that shows true or false a claim about the similarity of biometric reference(s)

and recognition biometric sample(s) by making a biometric comparison(s).

► Verification/Identification/Authentication/Recognition: The Terminology

Vetting

► Background Checks

Video Camera

► Face Device

Video Surveillance

► Human Detection and Tracking

Video-based Face Recognition

► Face Recognition, Video-based

Video-based Motion Capture

► Markerless 3D Human Motion Capture from Images

Visible Spectrum

Synonyms

Optical spectrum; Visible light

Definition

The portion of the electromagnetic spectrum that is visible (detected) by the human eye. The wavelengths for this spectrum is 380 to 750 nm, which are the wavelengths seen (detected) by the human eye in air.

► Iris Databases

Visual Memory

Visual memory is the perceptual ability that allows visual images to remain in memory after they are no longer visible. It supports the matching process between two fingerprints when eye movements are required.

► Latent Fingerprint Experts

Visual Sensor

► Face Device

Visual-dynamic Speaker Recognition

► Lip Movement Recognition

Vitality

► Liveness Detection: Fingerprint
► Liveness Detection: Iris

Viterbi Algorithm

The Viterbi algorithm is the conventional, recursive, efficient way to decode a Hidden Markov Model that is to find the optimal state sequence, given the observation sequence and the model. It provides information about the hidden process and is a good an efficient approximation of the evaluation problem.

► Hidden Markov Models

VOCs (Volatile Organic Compounds)

Organic chemicals that have a high vapor pressure resulting in a relatively high abundance in the headspace of samples.

► Odor Biometrics

Voice Authentication

Voice authentication is also known as speaker authentication, speaker verification, and one-to-one speaker recognition. For example, for a *client* – a bank customer – to be authenticated, the client must first go through an *enrollment* procedure, also known as training. During enrollment, the client provides a number of voice samples to the system, which in turn are used to build a *voice model* for the client. When requesting a voice authentication, a client must first announce his or her identity. This may be done verbally by saying name, user id, account number or the like, or it may be done by presenting an identifying token such as a staff card or bank card. Then the authentication takes place when the person speaks a set phrase or a requested phrase or simply engages in a dialogue with the authentication system. If the voice sample matches the stored model or template of the claimed identity, the client is authenticated. If an *impostor* tries to be authenticated as a particular client, the impostor's voice will not match the client model and the impostor will be rejected. The authentication paradigm only compares a speech sample with a single client model, namely the model of the claimed identity. Hence, it is sometimes known as one-to-one speaker recognition. In contrast *speaker identification* compares a speech sample with every possible client model, to find the closest match. Hence this paradigm is also known as one-to-many speaker recognition.

► Liveness Assurance in Voice Authentication
► Speaker Recognition Standardization

Voice Biometric

► Speaker Recognition, Overview

Voice Biometric Engine

► Speaker Matching

Voice Device

Doroteo T. Toledano,
Joaquin Gonzalez-Rodriguez,
Javier Ortega-Garcia
ATVS – Biometric Recognition Group, Escuela Politecnica Superior, Universidad Autonoma de Madrid, Spain

Synonyms

Microphone; Speech input device

Definition

Voice device in the context of biometrics is frequently used as a synonym for a simpler word: *microphone*. A microphone [1] is a transducer that converts sound (or equivalently, air pressure variations) into electrical signals. There are many different types of microphones that use different methods to achieve this transduction, most of which will be revised in this article. Besides the method employed to do the transduction, microphones are most frequently encapsulated, and the encapsulation allows to build microphones with different directional characteristics, which allow, for instance, to capture the voice coming from one direction and reject (to a certain extent) the noises or voices coming from other directions. Apart from the directionality, microphones also have different frequency responses, sensitivities, and dynamic ranges. All these characteristics can dramatically influence the performance of a speech biometric system, and should therefore be taken into account in the design of such systems.

Microphones are the most commonly used speech input devices, and for that reason they deserve most of the space of this article. However, this article will be incomplete without mentioning that microphones, at least traditional microphones, are not the only speech input device that can be used in speech biometrics. For instance, microphones may be arranged to form ▶ microphone arrays. There also exist special microphones called ▶ contact microphones that transduce vibrations in solid bodies into electrical signals. Finally, there is also the possibility of combining the acoustic evidence and the visual evidence of speech by recording the audio and also the movement of the lips in what is commonly referred to as audio-visual speech processing. Definitional entries are devoted at the end of this article to these special speech input devices.

The first step in any voice biometric (or automatic speaker recognition) system is to capture the voice of the speaker, and speech input devices are used for this purpose.

Introduction

The human hearing sense is extremely robust against noise and small distortions in the speech and humans are very good at recognizing people based on their voices, even under strong distortion and noise. Most speech input devices are designed with the goal of capturing speech or music, translating it into electrical signals, transmitting or storing it and, finally, reproducing that speech or music (by means of the opposite transducer, a *loudspeaker*). The important point here is that microphones are designed to be used in a chain, at which end is, most times, the human ear. Having such a robust receptor at the end of the chain makes it unnecessary to be very careful in the design or selection of a speech input device.

During the last years, however, there has been a fundamental change in speech communication since the receiver in the speech communication chain is not always a human listener any more. Nowadays machines are used for transcribing speech signals (in *automatic speech recognition*) and also, and most importantly in this context, for recognizing the speaker given a segment of speech (in *voice biometrics* or *automatic speaker recognition*). This fundamental change has brought an uncomfortable reality for all speech researchers: machines are still far less robust than humans at processing speech.

Of course, the goal of speech researchers is making machines not even as robust as humans but even more. Currently, voice biometric systems achieve very good results in relatively controlled conditions, such as in telephone conversations with similar durations. This has been the basic setup for the yearly competitive Speaker Recognition Evaluations (SRE) organized by the National Institute of Standards and Technology (NIST) [2] for the last years. These evaluations show that currently, technology is capable of achieving very competitive results in these conditions and is becoming more and more robust against variabilities.

However, the problem of variability due to the speech input device is far from being solved. Actually, this is a very hot research and technological topic. The proof of it is that next NIST evaluations in voice biometrics will probably be centered on *cross-channel* conditions in which training and testing data come from different channels (including different microphones, microphone locations (close-talk and far-field) and recording devices. However, achieving robustness against such variations is a long-term research goal that most probably will not be fulfilled in the next few years.

In the meantime, it should be stressed that technology is already usable in practical situations, but it should also be highlighted that current technology may not be as robust as desirable. In these circumstances it is essential to take extra care of the design or the selection of the speech input device. In some cases, of course, the speech input device is out of control, such as in telephone applications. But there are other cases where it is necessary to design the speech input device and, in this cases, it is essential to make the right choice because there are multiple choices of speech input devices with very different features, and an appropriate selection of the speech input device could be the key to success or failure in a voice biometrics application. This section tries to provide an introduction to the world of speech input devices or microphones.

Microphones

Definition

A microphone is a transducer that converts sounds (air pressure variations) into variations of an electrical magnitude, typically voltage.

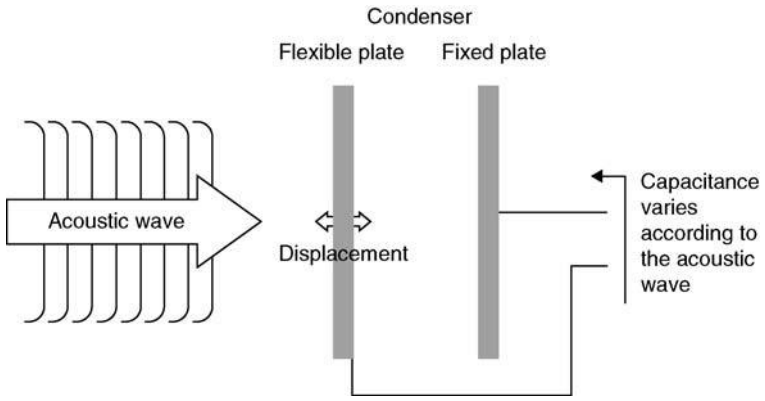
History

The early history of the microphone is tied to the development of the telephone [3]. In fact, the microphone was the last element required for a telephonic conversation to be developed. One of the earliest versions of microphones was developed by German researcher Philipp Reis in 1861. These microphones were just a platinum piece associated with a membrane that opened and closed an electric circuit as the sound made the membrane vibrate. This allowed Reis to build primitive prototypes that allowed to transmit voice and music along several hundred meters. It was several years later, in 1874, when Alexander Graham Bell patented the telephone and transmitted what is considered the first telephone conversation "*Mr. Watson, come here, I want you.*" Bell improved the microphones to make them better and better suited for commercial applications. Among the earlier microphones developed by Bell there are *liquid microphones* in which a diaphragm moved a metallic needle inside a metal recipient filled with a solution of water and sulfuric acid, so that the resistance between the needle and the recipient varied with the movement of the diaphragm. The latter microphones developed by Bell were based on the variations of inductance in a moving coil attached to a diaphragm. However, it was not until 1878 that the word *microphone* was used for the first time, and it was associated with what it is know today as the *carbon microphone*. The carbon microphone was invented by Edison and Hughes, and constituted a real breakthrough for telephone systems, since they were more efficient and robust than the earlier devices. Currently it has mostly been substituted by more modern microphones that will be described in the following sections.

Types

All microphones are based on the transduction of air pressure variations into an electromagnetic magnitude. However, there are many ways to achieve this, and therefore there are many types of microphones with different characteristics and applications. In this article some of the most important types will be summarized.

- *Condenser or capacitance microphones.* These microphones are based on the following physical principle (Fig. 1): the capacitance of a condenser with two metallic plates depends on the distance between the two plates. If one metallic plate of a capacitor is substituted by a metallic membrane that vibrates with sound, the capacitance of the condenser varies with sound, and this variation can be translated into the variation of an electrical magnitude. There are two ways of doing this transformation. The most common one is trying to set a constant charge in the two plates and measuring the variations of the voltage between the two plates.



Voice Device. **Figure 1** Principle of functioning of a condenser microphone.

The other one (slightly more complex) is using the variations in the capacitance to modulate the frequency of an oscillator. This generates a frequency modulated signal that needs to be demodulated, but the demodulated signal has usually less noise and can more effectively reproduce low frequency signals than the one obtained with the constant charge method. A special type of condenser microphone is the *electret microphone*. This microphone is a capacitor microphone in which the charge in the plates is maintained not by applying an external constant voltage to the capacitor, but by using a ferroelectric material that keeps a constant charge, in a similar way as a magnet generates a constant magnetic field. Condenser microphones are the most frequently used microphones nowadays, and it is possible to find them from low-quality cheap versions to high-quality expensive microphones.

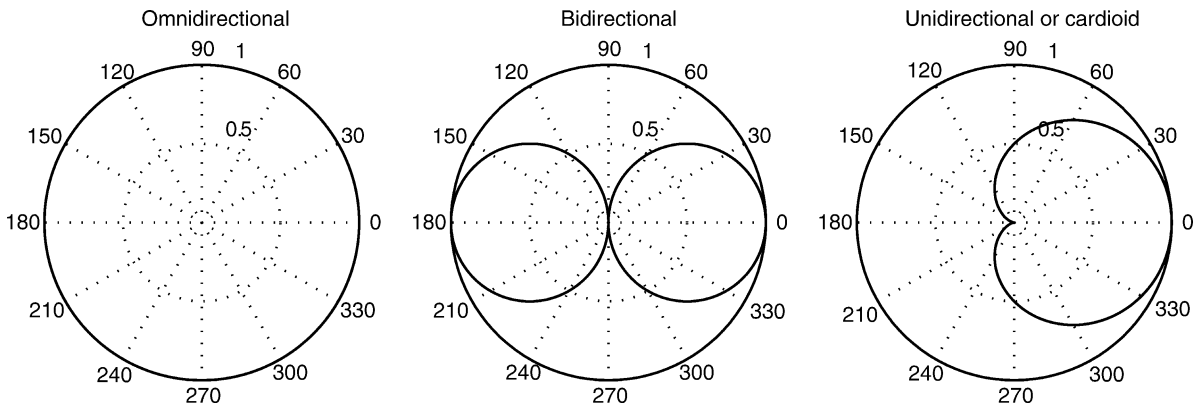
- *Dynamic or induction microphones.* These microphones are based on a different physical principle: when an induction moves inside a magnetic field, it generates a voltage by electromagnetic induction. If a small coil is attached to a diaphragm that moves with sounds and if this coil is placed into a magnetic field (generated by a permanent magnet), the movement of the coil will produce a voltage in its extremes that is related to the sound. A special type of induction microphone is *ribbon microphones* in which the coil is substituted by a metallic ribbon that vibrates with sound as is suspended in a constant magnetic field, thus generating a current related to the sound. These microphones are more sensitive than coil microphones, but also are more fragile.
- *Carbon microphones.* This microphone is essentially a recipient filled with carbon powder and closed by a metallic membrane on one side and a metallic plate on the other. As the membrane vibrates with the sound the powder supports more or less pressure and its electrical resistance is higher or lower (with more pressure carbon particles increase their surface in contact with other particles and this makes electrical resistance decrease). Carbon microphones were widely used in telephones. Currently they have been substituted by capacitor microphones.
- *Piezo-electric microphones.* These microphones are based on yet another physical effect: some solid materials, called *piezo-electric* materials, have the property of producing a voltage when a pressure is applied to them. Using this property and a piezo-electric material a microphone can by built by just placing two electrical contacts on the piezo-electric material. Piezo-electric microphones are mainly used in musical instruments (such as electric guitars) to collect and amplify the sound.
- *Silicon microphones.* Silicon (or chip) microphones are not based on a new physical effect. Rather, they are just capacitor microphones built on a silicon chip in which the membrane is directly attached to the chip. These microphones can be very small and are usually associated with electronic circuitry such as a preamplifier and a analog-to-digital converter (ADC), so that a single chip can produce digital audio.

Directional Characteristics

Microphones have different characteristics depending on the direction of arrival of the sound with respect to the microphone. A microphone's *directionality pattern* measures its sensitivity to a particular direction.

Microphones may be classified by their directional properties as *omnidirectional* (or *non-directional*) and *directional* [4]. The latter can also be subdivided into *bidirectional* and *unidirectional*, based on their directionality patterns. Directionality patterns are usually specified in terms of the *polar pattern* of the microphone (Fig. 2).

- *Omnidirectional microphones.* An omnidirectional (or nondirectional) microphone is a microphone whose response is independent of the direction of arrival of the sound wave. Sounds coming from different directions are picked equally. If a microphone is built only to respond to the pressure, then the resultant microphone is an omnidirectional microphone. These types of microphones are the most simple and inexpensive and have as advantage having a very flat frequency response. However, the property of capturing sounds coming from every direction with the same sensitivity is very often undesirable, since it is usually interesting capturing the sounds coming from the front of the microphone but not from behind or the laterals.
- *Bidirectional microphones.* If a microphone is built to respond to the gradient of the pressure in a particular direction, rather than to the pressure itself, a bidirectional microphone is obtained. This is achieved by letting the sound wave reach the diaphragm not only from the front of the microphone but also from the rear, so that if a wave comes from a perpendicular direction the effects on the front and the rear are canceled. This type of microphones reach the maximum sensitivity at the front and the rear, and reach their minimum sensitivity at the perpendicular directions. This directionality pattern is particularly interesting to reduce noises from the sides of the microphone. For this reason sometimes it is said that these microphones are *noise-canceling* microphones. Among the disadvantages of this kind of microphones, it must be mentioned that their frequency response is not nearly as flat as the one of an omnidirectional microphone, and it also varies with the direction of arrival. The frequency response is also different with the distance from the sound source to the microphone. Particularly, for sounds generated close to the microphone (near field) the response for low frequencies is higher than for sounds generated far from the microphone (far field). This is known as the *proximity effect*. For that reason frequency responses are given usually for far-field and near-field conditions, particularly for close-talking microphones. This type of microphones are more sensitive to the noises produced by the wind and the wind induced by the pronunciation of plosive sounds (such as /p/) in close-talking microphones.
- *Unidirectional Microphones.* These microphones have maximum response to sounds coming from the front of the microphone, have nearly zero response to sounds coming from the rear of the microphone and small response to sounds coming from the sides of the microphone. Unidirectionality is achieved by building a microphone that responds to the gradient of the sounds, similar to



Voice Device. Figure 2 Typical polar patterns for omnidirectional, bidirectional and unidirectional (or cardioid) microphones.

a bidirectional microphone. The null response from the rear is attained by introducing a material to slow down the acoustic waves coming from the rear so that when the wave comes from the rear it takes equal time to reach the rear part and the front part of the diaphragm, and therefore both cancel out. The polar pattern of these microphones has usually the shape of a heart, and for that reason are sometimes called *cardiod* microphones. These microphones have good noise-cancelation properties, and for these reasons, are very well suited for capturing clean audio input.

Microphone Location

Some microphones have different frequency response when the sound source is close to the microphone (near field, or close-talking) and when the sound source is far from the microphone (far field). In fact, not only the frequency response, but also the problems to the voice biometric application and the selection of the microphone could be different. For this reason a few concepts about microphone location will be reviewed.

- *Close-talking or near-field microphones.* These microphones are located close to the mouth of the speaker, usually pointing at the mouth of the speaker. This kind of microphones can benefit from the directionality pattern to capture mainly the sounds produced by the speaker, but could also be very sensitive to the winds produced by the speaker, if placed just in front of the mouth. The characteristics of the sound captured may be very different if the microphone is placed at different relative positions from the mouth, which is sometimes a problem for voice biometrics applications.
- *Far-field microphones.* These microphones are located at some distance from the speaker. They have the disadvantage that they tend to capture more noise than close-talking microphones because sometimes cannot take advantage of directionality patterns. This is particularly true if the speaker can move around as she speaks. In general, far-field microphone speech is considered to be far more difficult to process than close-talking speech. In some circumstances it is possible to take advantage of *microphone arrays* to locate the speaker spatially and to focus the array to listen specially to them.

Specifications

There is an international standard for microphone specifications [5], but few manufacturers follow it exactly. Among the most common specifications of a microphone the following must be mentioned.

- *Sensitivity.* The sensitivity measures the efficiency in the transduction (i.e. how much voltage it generates for an input acoustic pressure). It is measured in millivolts per Pascal at 1 kHz.
- *Frequency Response.* The frequency response is a measure of the variation of the sensitivity of a microphone as a function of the frequency of the signal. It is usually represented in decibels (dB) over a range of frequency typically between 0 and 20 kHz. The frequency response is dependent on the direction of arrival of the sound and the distance from the sound source. The frequency response is typically measured for sound sources very far from the microphone and with the sound reaching the microphone from its front direction. For close talking microphones it is also typical to represent the frequency response for sources close to the microphone to take into account the *proximity effect.*
- *Directional Characteristics.* The directionality of a microphone is the variation of its sensitivity as a function of the sound arrival direction, and is usually specified in the form of a directionality pattern, as explained earlier.
- *Non-Linearity.* Ideally, a microphone should be a linear transducer, and therefore a pure audio tone should produce a single pure voltage sinusoid at the same frequency. As microphones are not exactly linear, a pure acoustic tone produces a voltage sinusoid at the same frequency but also some harmonics. The most extended nonlinearity measure is the *total harmonic distortion, THD*, which is the ratio between the power of the harmonics produced and the power of the voltage sinusoid produced at the input frequency.
- *Limiting Characteristics.* These characteristics indicate the maximum sound pressure level (SPL) that can be transduced with limited distortion by the microphone. There are two different measures, the *maximum peak SPL* for a maximum THD, and the *overload, clipping or saturation level.* This last one indicates the SPL that produces the

maximum displacement of the diaphragm of the microphone.

- *Inherent Noise.* A microphone, in the absence of sound, produces a voltage level due to the inherent noise produced by itself. This noise is measured as the input SPL that would produce the same output voltage, which is termed the *equivalent SPL due to inherent noise.* This parameter determines the minimum SPL that can be effectively transduced by the microphone.
- *Dynamic Range.* The former parameters define the dynamic range of the microphone, (i.e. the minimum and maximum SPL that can be effectively transduced).

Summary

Speech input devices are the first element in a voice biometric system and are sometimes not given the attention they deserve in the design of voice biometric applications. This section has presented some of the variables to take into account in the selection or design of a microphone for a voice biometric application. The right selection, design, and even placement of a microphone could be crucial for the success of a voice biometric system.

Related Entries

- ▶ Biometric Sample Acquisition
- ▶ Sample Acquisition (System Design)
- ▶ Sensors

References

1. Eargle, J.: The Microphone Book, 2nd edn. Focal, Elsevier, Burlington, MA (2005)
2. National Institute of Standards and Technology (NIST): NIST Speaker Recognition Evaluation. http://www.nist.gov/speech/tests/spk/
3. Flichy, P.: Une Histoire de la Communication Moderne. La Decouverte (1997)
4. Huang, X., Acero, A., Hon, H.W.: Spoken Language Processing. Prentice-Hall PTR, New Jersey (2001)
5. International Electrotechnical Comission: International Standard IEC 60268-4: Sound systems equipment, Part 4: Microphones. Geneva, Switzerland (2004)

Voice Evidence

The forensic evidence of voice consists of the quantified degree of similarity between the speaker dependent features extracted from the questioned recording (trace) and the same extracted from recorded speech of a suspect, represented by his or her model.

▶ Voice, Forensic Evidence of

Voice Recognition

▶ Speaker Recognition, Overview
▶ Speaker Recognition Standardization

Voice Sample Synthesis

Juergen Schroeter
AT&T Labs – Research, Florham Park, NJ, US

Synonyms

Speech synthesis; Synthetic voice creation; Text-to-speech (TTS)

Definition

Over the last decade, speech synthesis, the technology that enables machines to talk to humans, has become so natural-sounding that a naïve listener might assume that he/she is listening to a recording of a live human speaker. Speech synthesis is not new; indeed, it took several decades to arrive where it is today. Originally starting from the idea of using physics-based models of the vocal-tract, it took many years of research to perfect the encapsulation of the acoustic properties of the vocal-tract as a "black box", using so-called formant synthesizers. Then, with the help of ever more

powerful computing technology, it became viable to use snippets of recorded speech directly and glue them together to create new sentences in the form of concatenative synthesizers. Combining this idea with now available methods for fast search, potentially millions of choices are evaluated to find the optimal sequence of speech snippets to render a given new sentence. It is the latter technology that is now prevalent in the highest quality speech synthesis systems. This essay gives a brief overview of the technology behind this progress and then focuses on the processes used in creating voice inventories for it, starting with recordings of a carefully-selected donor voice. The fear of abusing the technology is addressed by disclosing all important steps towards creating a high-quality synthetic voice. It is also made clear that even the best synthetic voices today still trip up often enough so as not to fool the critical listener.

Introduction

Speech synthesis is the technology that gives computers the ability to communicate to the users by voice. When driven by text input, speech synthesis is part of the more elaborate ► *text-to-speech (TTS) synthesis*, which also includes text processing (expanding abbreviations, for example), letter-to-sound transformation (rules, pronunciation dictionaries, etc.), and stress and pitch assignment [1]. Speech synthesis is often viewed as encompassing the signal-processing "backend" of text-to-speech synthesis viewed as encompassing the signal-processing "backend" of text-to-speech synthesis (with text and linguistic processing being carried out in the "front-end"). As such, speech synthesis takes *phoneme*-based information in context and transforms it into audible speech. Context information is very important because, in naturally-produced speech, no single speech sound stands by itself but is always highly influenced by what sounds came before, and what sounds will follow immediately after. It is precisely this context information that is key to achieving high-quality speech output.

A high-quality TTS system can be used for many applications, from telecommunications to personal use. In the telecom area, TTS is the only practical way to provide highly flexible speech output to the caller of an automated speech-enabled service. Examples of such services include reading back name and address information, and providing news or email reading. In the personal use area, the author has witnessed the ingenious "high jacking" of AT&T's web-based TTS demonstration by a young student to fake his mother's voice in a telephone call to his school: "Timmy will be out sick today. He cannot make it to school." It seems obvious that natural-sounding, high quality speech synthesis is vital for both kinds of applications. In the telecom area, the provider of an automated voice service might lose customers if the synthetic voice is unintelligible or sounds unnatural. If the young student wants to get an excused day off, creating a believable "real-sounding" voice seems essential. It is mostly concerns about the latter kind of potential abuse that motivates this author to write this essay. In the event that the even stricter requirement is added of making the synthetic voice indistinguishable from the voice of a specific person, there is clearly a significantly more difficult challenge. Shortly after AT&T's Natural Voices® TTS system became commercially available in August 2001, an article in the New York Times' Circuits section [2] asked precisely whether people will be safe from serious criminal abuse of this technology. Therefore, the purpose of this essay is to demystify the process of creating such a voice, disclose what processes are involved, and show current limitations of the technology that make it somewhat unlikely that speech synthesis could be criminally abused anytime soon.

This essay is organized as follows. The next section briefly summarizes different speech synthesis methods, followed by a somewhat deeper overview of the so-called Unit Selection synthesis method that currently delivers the highest quality speech output. The largest section of this essay deals with creating voice databases for unit selection synthesis. The essay concludes with an outlook.

Overview of Voice Synthesis Methods

The voice (speech) synthesis method with the most vision and potential, but also with somewhat unfulfilled promises, is articulatory synthesis. This method employs mathematical models of the speech production process in the human vocal tract, for example, models of the mechanical vibrations of the vocal chords (glottis) that interact with the fluid dynamics of the laminar and turbulent airflow from the lungs to the lips, plus linear or even nonlinear acoustical

models of sound generation and propagation along the vocal tract. A somewhat comprehensive review of this method is given in [3]. Due to high computational requirements and the need for highly accurate modeling, articulatory synthesis is mostly useful for research in speech production. It usually delivers unacceptably low-quality synthetic speech.

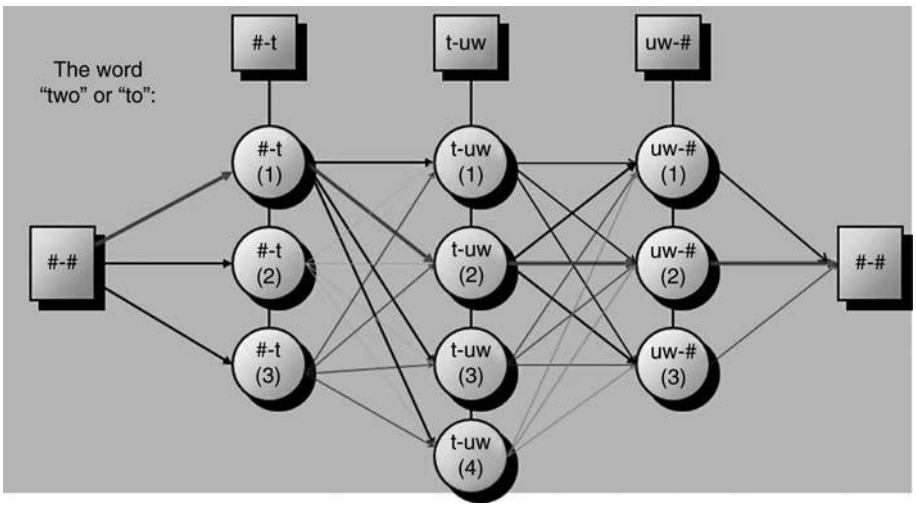
One level higher in abstraction, and much more practical in its use, is formant synthesis. This method captures the characteristics of the resonances of the human vocal tract in terms of simple filters. The single-peaked frequency characteristic of such a filter element is called formant. Its frequency, bandwidth (narrow to broad), and amplitude fully specify each formant. For adult vocal tracts, four to five formants are enough to determine their acoustic filter characteristics. Phonetically most relevant are the lowest three formants that span the vowel and sonorant space of a speaker and a language. Together with a suitable waveform generator that approximates the glottal pulse, formant synthesis systems, due to their highly versatile control parameter sets, are very useful for speech perception research. More on formant synthesis can be found in [4]. For use as a speech synthesizer, the computational requirements are relatively low, making this method the preferred option for embedded applications, such as reading back names (e.g., "calling Mom") in a dial-by-voice cellular phone handset. Its storage requirements are miniscule (as little as 1 MB). Formant synthesis delivers intelligible speech when special care is given to consonants.

In the 1970s, a new method started to compete with the, by then, well-established formant synthesis method. Due to its main feature of stitching together recorded snippets of natural speech, it was called concatenative synthesis. Many different options exist for selecting the specific kind of elementary speech units to concatenate. Using words as such units, although intuitive, is not a good choice given that there are many tens of thousands of them in a language and that each recorded word would have to fit into several different contexts with its neighbors, creating the need to record several versions of each word. Therefore, word-based concatenation usually sounds very choppy and artificial. However, subword units, such as *diphones* or *demisyllables* turned out to be much more useful because of favorable statistics. For English, there is a minimum of about 1500 ▶ diphones that would need to be in the inventory of a diphone-based concatenative synthesizer. The number is only slightly higher for concatenating ▶ demisyllables. For both kinds of units, however, elaborate methods are needed to identify the best single (or few) instances of units to store in the voice inventory, based on statistical measures of acoustic typicality and ease of concatenation, with a minimum of audible glitches. In addition, at synthesis time, elaborate speech signal processing is needed to assure smooth transitions, deliver the desired *prosody*, etc. For more details on this method, see [5]. Concatenative synthesis, like formant synthesis, delivers highly intelligible speech and usually has no problem with transients like stop consonants, but usually lacks naturalness and thus cannot match the quality of direct human voice recordings. Its storage requirements are moderate by today's standards (∼10–100 MB).

Unit Selection Synthesis

The effort and care given to creating the voice inventory determines to a large extent the quality of any concatenative synthesizer. For best results, most concatenative synthesis researchers well up into the 1990s employed a largely manual off-line process of trial and error that relied on dedicated experts. A selected unit needed to fit all possible contexts (or made to fit by signal processing such as, stretching or shrinking durations, pitch scaling, etc.). However, morphing any given unit by signal processing in the synthesizer at synthesis time degrades voice quality. So, the idea was born to minimize the use of signal processing by taking advantage of the ever increasing power of computers to handle ever increasing data sets. Instead of outright morphing a unit to make it fit, the synthesizer may try to pick a suitable unit from a large number of available candidates, optionally followed by much more moderate signal processing. The objective is to find automatically the optimal sequence of unit instances at synthesis time, given a large inventory of unit candidates and the available sentence to be synthesized. This new objective turned the speech synthesis problem into a rapid search problem [6].

The process of selecting the right units in the inventory that instantiate a given input text, appropriately called unit selection, is outlined in Fig. 1. Here, the word "two" (or "to") is synthesized from using diphone candidates for silence into "t" (/#-t/), "t" into "uw" (/t-uw/), and "uw" into silence (/uw-#/).



Voice Sample Synthesis. **Figure 1** Viterbi search to retrieve optimal diphone units for the word "two" or "to".

Each time slot (column in Fig. 1) has several candidates to choose from. Two different objective distance measures are employed. First, transitions from one unit to the next (depicted by arrows in the figure) are evaluated by comparing the speech spectra at the end of the left-side unit candidates to the speech spectra at the beginning of the right-side unit candidates. These are n^*m comparisons, where n is the number of unit candidates for the left column of candidates, and m is the number of unit candidates in the right-side column of candidates. Second, each node (circle) in the network of choices depicted in Fig. 1 has an intrinsic "goodness of fit" measured by a so-called target cost. The ideal target cost of a candidate unit measures the acoustic distance of the unit against a hypothetical unit cut from a perfect recording of the sentence to be synthesized. However, since it is unlikely that the exact sentence would be in the inventory, an algorithm has to estimate the target cost using symbolic and nonacoustic cost components such as the difference between desired and given pitch, amplitude, and context (i.e., left and right phone sequences).

The objective of selecting the optimal unit sequence for a given sentence is to minimize the total cost that is accumulated by summing transitional and target costs for a given path through the network from its left-side beginning to its right-side end. The optimal path is the one with the minimum total cost. This path can be identified efficiently using the Viterbi search algorithm [7].

More detailed information about unit selection synthesis can be found in [1, 8]. The latter book chapter also summarizes the latest use of automatic speech recognition (ASR) technology in unit selection synthesis.

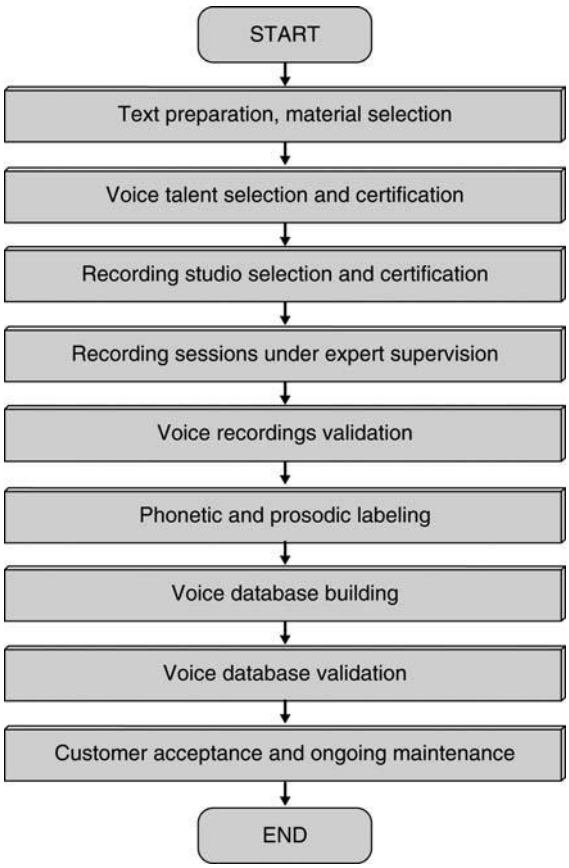
Voice Creation

Creating a simple-minded unit selection synthesizer would involve just two steps: First, record *exactly* the sentences that a user wants the machine to speak; and second, identify at "synthesis" time the input sentence to be spoken, and then play it back. In practice units are used that are much shorter than sentences to be able to create previously unseen input sentences, so this simple-minded paradigm would not work. However, when employing a TTS front-end that converts any input text into a sequence of unit specifications, intuition may ask for actually playing back any inventory sentence in its entirety in the odd chance that the corresponding text has been entered. Since the translation of text into unit-based tags and back into speech is not perfect, the objective is unlikely to ever be fully met. In practice, however, the following, somewhat weaker objective holds: as long as the text to be synthesized is similar enough to that of a corresponding recording that actually exists in the inventory, a high output voice quality can be expected. It is for this reason that unit-selection synthesis is particularly well suited for so-called limited domain synthesis, such as

weather reports, stock reports, or any automated telecom dialogue application (banking, medical, etc.) where the application designer can afford the luxury of recording a special inventory, using a carefully selected voice talent. High quality synthesis for general news or email reading is usually much more difficult to achieve because of coverage issues [9].

Because unit selection synthesis, to achieve its best quality results, mimics a simple tape recorder playback, it is obvious that its output voice quality largely depends on what material is in its voice inventory. Without major modifications/morphing at synthesis time, the synthesizer output is confined to the quality, speaking style, and emotional state of the voice that was recorded from the voice talent/donor speaker. For this reason, careful planning of the voice inventory is required. For example, if the inventory contains only speech recorded from a news anchor, the synthesizer will always sound like a news anchor.

Several issues need to be addressed in planning a voice inventory for a unit selection synthesizer. The steps involved are outlined in Fig. 2, starting with text preparation to cover the material selected. Since voice recordings cannot be done faster than real time, they are always a major effort in time and expense. To get optimal results, a very strict quality assurance process for the recordings is paramount. Furthermore, the content of the material to be recorded needs to be addressed. Limited domain synthesis covers typical text for the given application domain, including greetings, apologies, core transactions, and good-byes. For more general use such as email and news reading, potentially hundreds of hours of speech need to be recorded. However, the base corpus for both kinds of applications needs to maximize linguistic coverage within a small size. Including a core corpus that was optimized for traditional diphone synthesis might satisfy this need. In addition, news material, sentences that use the most common names in different prosodic contexts, addresses, and greetings are useful. For limited domain applications, domain-specific scripts need to be created. Most of them require customer input such as getting access to text for existing voice prompts, call flows, etc. There is a significant danger in underestimating this step in the planning phase. Finally, note that a smart and frugal effort in designing the proper text corpus to record helps to reduce the amount of data to be recorded. This, in turn, will speed up the rest of the voice building process.



Voice Sample Synthesis. **Figure 2** Steps in unit selection voice inventory creation.

Quality assurance starts with selecting the best professional voice talent. Besides the obvious criteria of voice preference, accent, pleasantness, and suitability for the task (a British butler voice might not be appropriate for reading instant messages from a banking application), the voice talents needs to be very consistent in how she/he pronounces the same word over time and in different contexts. Speech production issues might come into play, such as breath noise, frequent lip smacks, disfluencies, and other speech defects. A clearly articulated and pleasant sounding voice and a natural prosodic quality are important. The same is true for consistency in speaking rate, level, and style. Surprisingly, good sight reading skills are not very common among potential voice talents. Speakers with heavy vocal fry (glottal vibration irregularities) or strong nasality should be avoided. Overall, a low ratio of usable recordings to total recordings done in a test run is a good criterion for rejecting a voice talent.

Pronunciations of rare words, such as foreign names, need to be agreed upon beforehand and their realizations monitored carefully. Therefore, phonetic supervision has to be part of all recording sessions.

Next, the recording studio used for the recording sessions should have almost "anechoic" acoustic characteristics and a very low background noise in order to avoid coloring or tainting the speech spectrum in any way. Since early acoustic reflections off a nearby wall or table are highly dependent on the time-varying geometry relative to the speaker's mouth and to the microphone, the recording engineer needs to make sure that the speaker does not move at all (unrealistic) or minimize these reflections. The recording engineer also needs to make sure that sound levels, and trivial things like the file format of the recordings are consistent and on target. Finally, any recorded voice data needs to be validated and inconsistencies between desired text and actually spoken text reconciled (e.g., the speaker reads "vegetarian" where "veterinarian" was requested).

Automatic labeling of large speech corpora is a crucial step because manual labeling by linguists is slow (up to 500 times real time) and potentially inconsistent (different human labelers disagree). Therefore, an automatic speech recognizer (ASR) is used in so-called forced alignment mode for phonetic labeling. Given the text of a sentence, the ASR identifies the identities and the beginnings and ends of all ▶ phonemes. ASR might employ several passes, starting from speaker-independent models, and adapting these models to the given single speaker, and his/her speaking style. Adapting the pronunciation dictionary to the specific speaker's individual pronunciations is vital to get the correct phoneme sequence for each recorded word. Pronunciation dictionaries used for phonetic labeling should also be used in the synthesizer. In addition, an automated *prosodic* labeler is useful for identifying typical stress and pitch patterns, prominent words, and phrase boundaries. Both kinds of automatic labeling need to use paradigms and conventions (such as phoneme sets and symbolic ▶ prosody tags) that match those used in the TTS front-end at synthesis time. A good set of automatic labeling and other tools allowed the author's group of researchers to speed up their voice building process by more than 100 times over 6 years.

Once the recordings are done, the first step in the voice building process is to build an index of which sound (phoneme) is where, normalize the amplitudes, and extract acoustic and segmental features, and then build distance tables used to trade off (weigh) different cost components in unit selection in the last section. One important part of the runtime synthesizer, the so-called Unit Preselection (a step used to narrow down the potentially very large number of candidates) can be sped up by looking at statistics of triples of phonemes (i.e., so-called triphones) and caching the results. Then, running a large independent training text corpus through the synthesizer and gathering statistics of unit use can be used to build a so-called join cache that eliminates recomputing join costs at runtime for a significant speedup. The final assembly of the voice database may include reordering of units for access efficiency plus packaging the voice data and indices.

Voice database validation consists of comprehensive, iterative testing with the goal of identifying bad units, either by automatic identification tools or by many hours of careful listening and "detective" work (where did this bad sound come from?), plus repair. Allocating sufficient testing time before compute-intensive parts of the voice building process (e.g., cache building) is a good idea. Also, setting realistic expectations with the customer (buyer of the voice database) is vital. For example, the author found that the "damage" that the TTS-voice creation and synthesis process introduces relative to a direct recording seems to be somewhat independent of the voice talent. Therefore, starting out with a "bad" voice talent will only lead to a poorer sounding synthetic voice. Reducing the TTS damage over time is the subject of ongoing research in synthesis-related algorithms employed in voice synthesis.

The final step in unit selection voice creation is formal customer acceptance and, potentially, ongoing maintenance. Formal customer acceptance is needed to avoid disagreements over expected and delivered quality, coverage, etc. Ongoing maintenance assures high quality for slightly different applications or application domains, including, for example, additional recordings.

Conclusion

This essay highlighted the steps involved in creating a high-quality sample-based speech synthesizer. Special focus was given to the process of voice inventory creation.

From the details in this essay, it should be clear that voice inventory creation is not trivial. It involves many weeks of expert work and, most importantly, full collaboration with the chosen voice talent. The idea of (secretly) recording any person and creating a synthetic voice that sounds just like her or him is simply impossible, given the present state of the art. Collecting several hundreds of hours of recordings necessary to having a good chance at success of creating such a voice inventory is only practical when high-quality archived recordings are already available that were recorded under very consistent acoustic conditions. A possible workable example would be an archive containing a year or more of evening news read by a well-known news anchor. Even then, however, one would need to be concerned about voice consistency, since even slight cold infections, as well as more gradual natural changes over time (i.e., caused by aging of the speaker) can make such recordings unusable.

An interesting extension to the sample synthesis of (talking) faces was made in [10]. The resulting head-and-shoulder videos of synthetic personal agents are largely indistinguishable from video recordings of the face talent. Again, similar potential abuse issues are a concern.

One specific concern is that unit-selection voice synthesis may "fool" automatic speaker verification systems. Unlike a human listener's ear that is able to pick up the subtle flaws and repetitiveness of a machine's renderings of a human voice, today's speaker verification systems are not (yet) designed to pay attention to small blurbs and glitches that are a clear giveaway of a unit selection synthesizer's output, but this could change if it became a significant problem. If this happens, perceptually undetectable watermarking is an option to identify a voice (or talking face) sample as "synthetic". Other procedural options include asking for a second rendition of the passphrase and comparing the two versions. If they are too similar (or even identical), reject the speaker identity claim as bogus.

Related Entries

- ▶ Hidden Markov Model (HMM)
- ▶ Speaker Databases and Evaluation
- ▶ Speaker Matching
- ▶ Speaker Recognition, Overview
- ▶ Speech Production

References

1. Schroeter, J.: Basic principles of speech synthesis, In: Benesty, J. (ed.) Springer Handbook of Speech Processing and Communication, Chap. 19 (2008)
2. Bader, J.L.: Presidents as pitchmen, and posthumous play-by-play, commentary in the New York Times, August 9 (2001)
3. van Santen, J., Sproat, R., Olive, J., Hirschberg, J., (eds.): Progress in speech synthesis, section III. Springer, NY (1997)
4. Holmes, J.N.: Research report formant synthesizers: cascade or parallel? Speech Commun. **2**(4), 251–273 (1983)
5. Sproat, R. (ed.): Multilingual text-to-speech synthesis. The bell labs approach. Kluwer Academic Publishers, Dordrecht MA (1998)
6. Hunt, A., Black, A.W.: Unit selection in a concatenative speech synthesis system using a large speech database. In: Proceedings of the ICASSP-96, pp. 373–376, GA, USA (1996)
7. Forney, G.D.: The viterbi algorithm. Proc. IEEE **61**(3), 268–278 (1973)
8. Dutoit, T.: Corpus-based speech synthesis, In: Benesty, J. (ed.) Springer Handbook of Speech Processing and Communication, Chap. 21 (2008)
9. van Santen, J.: Prosodic processing. In: Benesty, J. (ed.) Springer Handbook of Speech Processing and Communication, Chap. 23 (2008)
10. Cosatto, E., Graf, H.P., Ostermann, J., Schroeter, J.: From audio-only to audio and video text-to-speech. Acta Acustica **90**, 1084–1095 (2004)

Voice Verification

▶ Liveness Assurance in Voice Authentication

Voice, Forensic Evidence of

Andrzej Drygajlo
Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland

Synonym

Forensic speaker recognition

Definition

Forensic speaker recognition is the process of determining if a specific individual (suspected speaker) is the

source of a questioned voice recording (trace). The forensic application of speaker recognition technology is one of the most controversial issues within the wide community of researchers, experts, and police workers. This is mainly due to the fact that very different methods are applied in this area by phoneticians, engineers, lawyers, psychologists, and investigators. The approaches commonly used for speaker recognition by forensic experts include the aural-perceptual, the auditory-instrumental, and the automatic methods. The forensic expert's role is to testify to the worth of the evidence by using, if possible a quantitative measure of this worth. It is up to other people (the judge and/or the jury) to use this information as an aid to their deliberations and decision.

This essay aims at presenting forensic automatic speaker recognition (FASR) methods that provide a coherent way of quantifying and presenting recorded voice as scientific evidence. In such methods, the evidence consists of the quantified degree of similarity between speaker-dependent features extracted from the trace and speaker-dependent features extracted from recorded speech of a suspect. The interpretation of a recorded voice as evidence in the forensic context presents particular challenges, including within-speaker (within-source) variability, between-speakers (between-sources) variability, and differences in recording sessions conditions. Consequently, FASR methods must provide a probabilistic evaluation which gives the court an indication of the strength of the evidence given the estimated within-source, between-sources, and between-session variabilities.

Introduction

Speaker recognition is the general term used to include all of the many different tasks of discriminating people based on the sound of their voices. Forensic speaker recognition involves the comparison of recordings of an unknown voice (questioned recording) with one or more recordings of a known voice (voice of the suspected speaker) [1, 2].

There are several types of forensic speaker recognition [3, 4]. When the recognition employs any trained skill or any technologically-supported procedure, the term technical forensic speaker recognition is often used. In contrast to this, so-called naïve forensic speaker recognition refers to the application of un-reflected everyday abilities of people to recognize familiar voices.

The approaches commonly used for technical forensic speaker recognition include the aural-perceptual, auditory-instrumental, and automatic methods [2]. Aural-perceptual methods, based on human auditory perception, rely on the careful listening of recordings by trained phoneticians, where the perceived differences in the speech samples are used to estimate the extent of similarity between voices [3]. The use of aural-spectrographic speaker recognition can be considered as another method in this approach. The exclusively visual comparison of spectrograms in what has been called the "▶ voiceprint " approach has come under considerable criticism in the recent years [5]. The auditory-instrumental methods involve the acoustic measurements of various parameters, such as the average fundamental frequency, articulation rate, formant centre-frequencies, etc. [4]. The means and variances of these parameters are compared. FASR is an established term used when automatic speaker recognition methods are adapted to forensic applications. In automatic speaker recognition, the statistical or deterministic models of acoustic features of the speaker's voice and the acoustic features of questioned recordings are compared [6].

FASR offers data-driven methodology for quantitative interpretation of recorded speech as evidence. It is a relatively recent application of digital speech signal processing and pattern recognition for judicial purposes and particularly law enforcement. Results of FASR based investigations may be of pivotal importance at any stage of the course of justice, be it the very first police investigation or a court trial. FASR has been gaining more and more importance ever since the telephone has become an almost ideal tool for the commission of certain criminal offences, especially drug dealing, extortion, sexual harassment, and hoax calling. To a certain degree, this is undoubtedly a consequence of the highly-developed and fully automated telephone networks, which may safeguard a perpetrator's anonymity. Nowadays, speech communications technology is accessible anywhere, anytime and at a low price. It helps to connect people, but unfortunately also makes criminal activities easier. Therefore, the identity of a speaker and the interpretation of recorded speech as evidence in the forensic context are quite often at issue in court cases [1, 7].

Although several speaker recognition systems for commercial applications (mostly speaker verification) have been developed over the past 30 years, until recently the development of a reliable technique for FASR has been unsuccessful because methodological aspects concerning automatic recognition of speakers in criminalistics and the role of the forensic expert have not been investigated sufficiently [8]. The role of a forensic expert is to testify in court using, if possible, quantitative measures that estimate the value and strength of the evidence. The judge and/or the jury use the testimony as an aid to the deliberations and decisions [9].

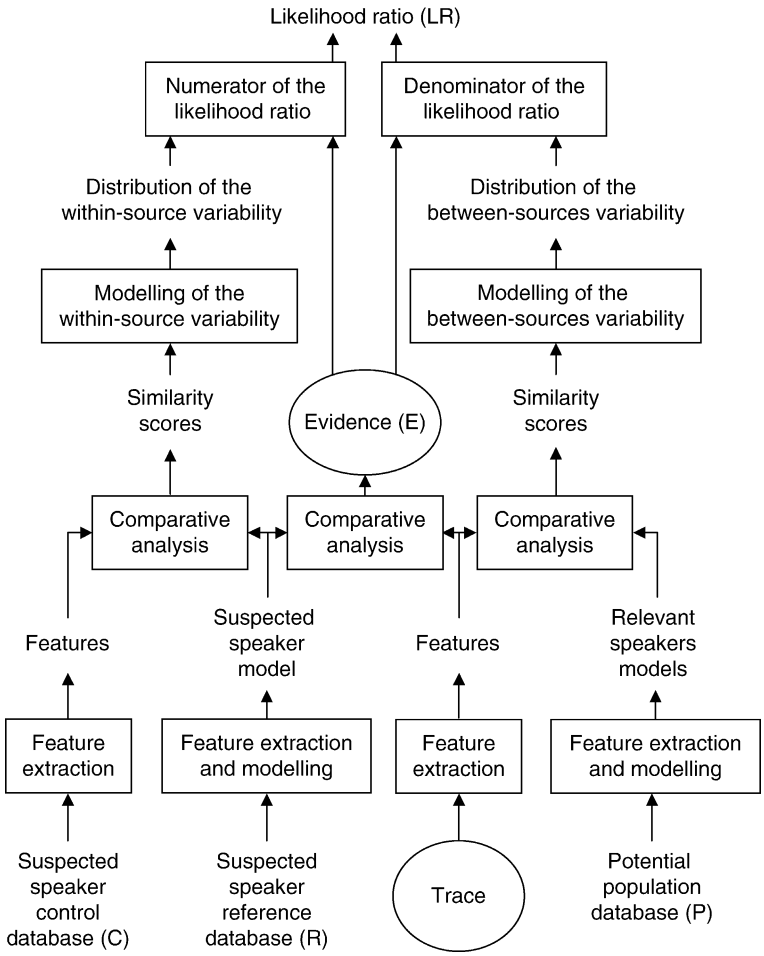
A forensic expert testifying in court is not an advocate, but a witness who presents factual information and offers a professional opinion based upon that factual information. In order for it to be effective, it must be carefully documented, and expressed with precision in neutral and objective way with the adversary system in mind. Technical concepts based on digital signal processing and pattern recognition must be articulated in layman terms such that the judge and the attorneys may understand them. They should also be developed according to specific recommendations that take into account also the forensic, legal, judicial, and criminal policy perspectives. Therefore, forensic speaker recognition methods should be developed based on current state-of-the-art interpretation of forensic evidence, the concept of identity used in criminalistics, a clear understanding of the inferential process of identity, and the respective duties of the actors involved in the judicial process, jurists, and forensic experts.

Voice as Evidence

When using FASR, the goal is to identify whether an unknown voice of a questioned recording (trace) belongs to a suspected speaker (source). The ► voice evidence consists of the quantified degree of similarity between speaker dependent features extracted from the trace, and speaker dependent features extracted from recorded speech of a suspect, represented by his or her model [1], so the evidence does not consist of the speech itself. To compute the evidence, the processing chain illustrated in Fig. 1 may be employed [10]. As a result, the suspect's voice can be recognized as the recorded voice of the trace, to the extent that the evidence supports the hypothesis that the questioned and the suspect's recorded voices were generated by the same person (source) rather than the hypothesis that they were not. However, the calculated value of evidence does not allow the forensic expert alone to make an inference on the identity of the speaker.

As no ultimate set of speaker specific features is present or detected in speech, the recognition process remains in essence a statistical-probabilistic process based on models of speakers and collected data, which depend on a large number of design decisions. Information available from the auditory features and their evidentiary value depend on the speech organs and language used [3]. The various speech organs have to be flexible to carry out their primary functions such as eating and breathing as well as their secondary function of speech, and the number and flexibility of the speech organs results in a high number of "degrees of freedom" when producing speech. These "degrees of freedom" may be manipulated at will or may be subject to variation due to external factors such as stress, fatigue, health, and so on. The result of this plasticity of the vocal organs is that no two utterances from the same individual are ever identical in a physical sense. In addition to this, the linguistic mechanism (language) driving the vocal mechanism is itself far from invariant. We are all aware of changing the way we speak, including the loudness, pitch, emphasis, and rate of our utterances; aware, probably, too, that style, pronunciation, and to some extent dialect, vary as we speak in different circumstances. Speaker recognition thus involves a situation where neither the physical basis of a person's speech (the vocal organs) nor the language driving it, are constant.

The speech signal can be represented by a sequence of short-term feature vectors. This is known as feature extraction (Fig. 1). It is typical to use features based on the various speech production and perception models. Although there are no exclusive features conveying speaker identity in the speech signal, from the source-filter theory of speech production it is known that the speech spectrum envelope encodes information about the speaker's vocal tract shape [11]. Thus some form of spectral envelope based features is used in most speaker recognition systems even if they are dependent on external recording conditions. Recently, the majority of speaker recognition systems have converged to the use of cepstral features derived from the envelope spectra models [1].



Voice, Forensic Evidence of. **Figure 1** Block diagram of the evidence processing and interpretation system. © IEEE.

Thus, the most persistent real-world challenge in this field is the variability of speech. There is within-speaker (within-source) variability as well as between-speakers (between-sources) variability. Consequently, forensic speaker recognition methods should provide a statistical-probabilistic evaluation, which attempts to give the court an indication of the strength of the evidence, given the estimated within-source variability and the between-sources variability [4, 10].

Bayesian Interpretation of Evidence

To address these variabilities, a probabilistic model [9], Bayesian inference [8] and data-driven approaches [6] appear to be adequate: in FASR statistical techniques the distribution of various features extracted from a suspect's speech is compared with the distribution of the same features in a reference population with respect to the questioned recording. The goal is to infer the identity of a source [9], since it cannot be known with certainty.

The inference of identity can be seen as a reduction process, from an initial population to a restricted class, or, ultimately, to unity [8]. Recently, an investigation concerning the inference of identity in forensic speaker recognition has shown the inadequacy of the speaker verification and speaker identification (in closed set and in open set) techniques [8]. Speaker verification and identification are the two main automatic techniques of speech recognition used in commercial applications. When they are used for forensic speaker recognition they imply a final discrimination decision based on a threshold. Speaker verification is the task of

deciding, given a sample of speech, whether a specified speaker is the source of it. Speaker identification is the task of deciding, given a sample of speech, which among many speakers is the source of it. Therefore, these techniques are clearly inadequate for forensic purposes, because they force the forensic expert to make decisions which are devolved upon the court. Consequently, the state-of-the-art speaker recognition algorithms using dynamic time warping (DTW) and hidden Markov models (HMMs) for text-dependent tasks, and vector quantization (VQ), Gaussian mixture models (GMMs), ergodic HMMs and others for text-independent tasks have to be adapted to the Bayesian interpretation framework which represents an adequate solution for the interpretation of the evidence in the judicial process [9].

The court is faced with decision-making under uncertainty. In a case involving FASR it wants to know how likely it is that the speech samples of questioned recording have come from the suspected speaker. The answer to this question can be given using the Bayes' theorem and a data-driven approach to interpret the evidence [1, 7, 10].

The odds form of Bayes' theorem shows how new data (questioned recording) can be combined with prior background knowledge (prior odds (province of the court)) to give posterior odds (province of the court) for judicial outcomes or issues (Eq. 1). It allows for revision based on new information of a measure of uncertainty (likelihood ratio of the evidence (province of the forensic expert)) which is applied to the pair of competing hypotheses: H_0 – the suspected speaker is the source of the questioned recording, H_1 – the speaker at the origin of the questioned recording is not the suspected speaker.

$$\underbrace{\frac{p(H_0|E)}{p(H_1|E)}}_{\substack{\text{posterior knowledge}\\ \textit{posterior odds}\\ \textit{(province of the court)}}} = \underbrace{\frac{p(E|H_0)}{p(E|H_1)}}_{\substack{\text{new data}\\ \textit{likelihood ratio}\\ \textit{(province of the expert)}}} \cdot \underbrace{\frac{p(H_0)}{p(H_1)}}_{\substack{\text{prior knowledge}\\ \textit{prior odds}\\ \textit{(province of the court)}}} \qquad (1)$$

This hypothetical-deductive reasoning method, based on the odds form of the Bayes' theorem, allows evaluating the likelihood ratio of the evidence that leads to the statement of the degree of support for one hypothesis against the other. The ultimate question relies on the evaluation of the probative strength of this evidence provided by an automatic speaker recognition method [12]. Recently, it was demonstrated that outcome of the aural (subjective) and instrumental (objective) approaches can also be expressed as a Bayesian likelihood ratio [4, 13].

Strength of Evidence

The ▶ strength of voice evidence is the result of the interpretation of the evidence, expressed in terms of the likelihood ratio of two alternative hypotheses. The principal structure for the calculation and the interpretation of the evidence is presented in Fig. 1. It includes the collection (or selection) of the databases, the automatic speaker recognition and the Bayesian interpretation [10].

The methodological approach based on a Bayesian interpretation (BI) framework is independent of the automatic speaker recognition method chosen, but the practical solution presented in this essay as an example uses text-independent speaker recognition system based on Gaussian mixture model (GMM) [14].

The Bayesian interpretation (BI) methodology needs a two-stage statistical approach [10]. The first stage consists in modeling multivariate feature data using GMMs. The second stage transforms the data to a univariate projection based on modeling the similarity scores. The exclusively multivariate approach is also possible but it is more difficult to articulate in layman terms [15]. The GMM method is not only used to calculate the evidence by comparing the questioned recording (trace) to the GMM of the suspected speaker (source), but it is also used to produce data necessary to model the within-source variability of the suspected speaker and the between-sources variability of the potential population of relevant speakers, given the questioned recording. The interpretation of the evidence consists of calculating the likelihood ratio using the probability density functions (pdfs) of the variabilities and the numerical value of evidence.

The information provided by the analysis of the questioned recording (trace) leads to specify the initial reference population of relevant speakers (potential population) having voices similar to the trace, and,

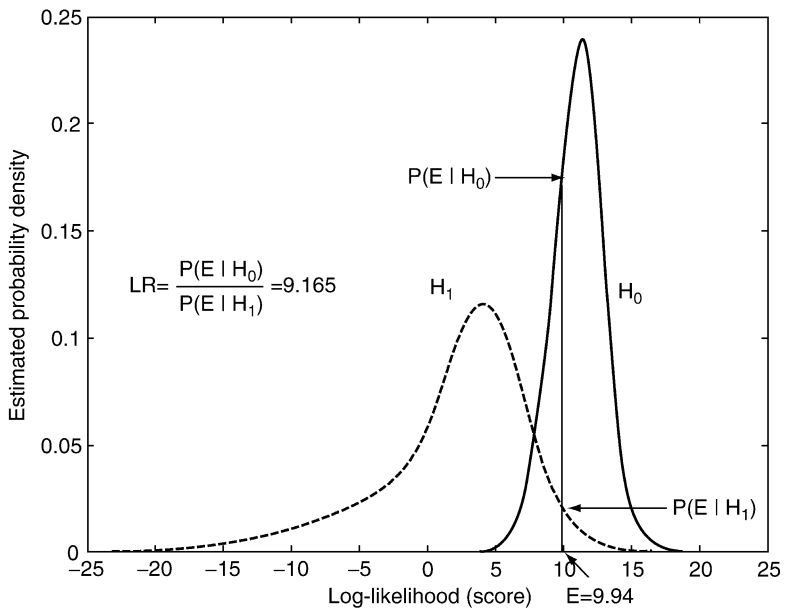
combined with the police investigation, to focus on and select a suspected speaker. The methodology presented needs three databases for the calculation and the interpretation of the evidence: the potential population database (P), the suspected speaker reference database (R), and the suspected speaker control database (C) [14].

The potential population database (P) is a database for modeling the variability of the speech of all the potential relevant sources, using the automatic speaker recognition method. It allows evaluating the between-sources variability given the questioned recording, which means the distribution of the similarity scores that can be obtained, when the questioned recording is compared to the speaker models (GMMs) of the potential population database. The calculated between-sources variability pdf is then used to estimate the denominator of the likelihood ratio $p(E|H_1)$. Ideally, the technical characteristics of the recordings (e.g., signal acquisition and transmission) should be chosen according to the characteristics analyzed in the trace.

The suspected speaker reference database (R) is recorded with the suspected speaker to model his/her speech with the automatic speaker recognition method. In this case, speech utterances should be produced in the same way as those of the P database. The suspected speaker model obtained is used to calculate the value of the evidence, by comparing the questioned recording to the model.

The suspected speaker control database (C) is recorded with the suspected speaker to evaluate her/his within-source variability, when the utterances of this database are compared to the suspected speaker model (GMM). This calculated within-source variability pdf is then used to estimate the numerator of the likelihood ratio $p(E|H_0)$. The recording of the C database should be constituted of utterances as far as possible equivalent to the trace, according to the technical characteristics, as well as to the quantity and style of speech.

The basic method proposed has been exhaustively tested in mock forensic cases corresponding to real caseworks [11, 14]. In an example presented in Fig. 2, the strength of evidence, expressed in terms of likelihood ratio gives $LR = 9.165$ for the evidence value $E = 9.94$, in this case. This means that it is 9.165 times more likely to observe the score E given the hypothesis H_0 than H_1. The important point to be made here is that the estimate of the LR is only as good as the modeling techniques and databases used to derive it. In the example, the GMM technique was used to estimate pdfs from the data representing similarity scores [11].



Voice, Forensic Evidence of. **Figure 2** The LR estimation given the value of the evidence *E*. © IEEE.

Evaluation of the Strength of Evidence

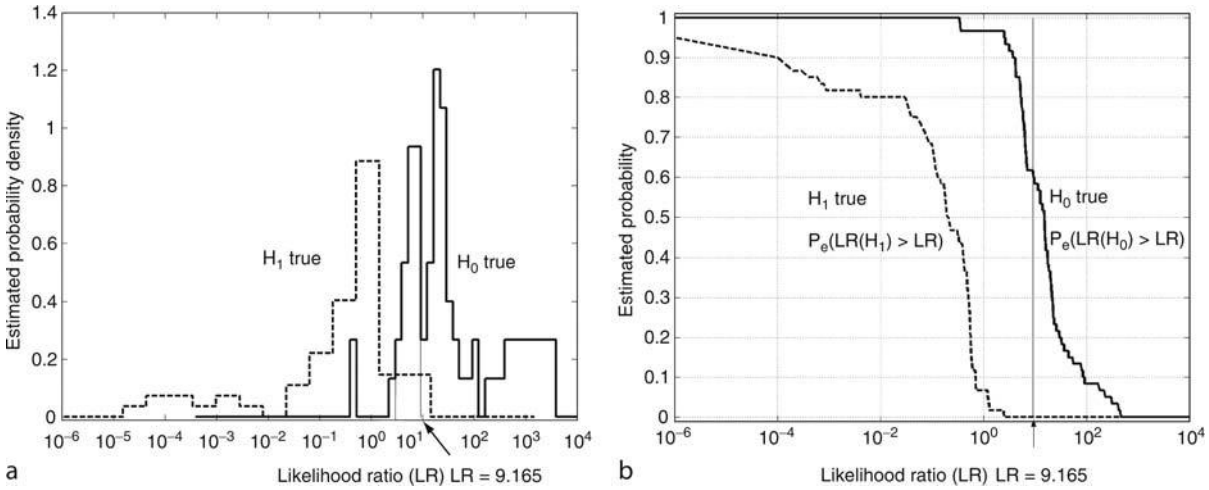
The likelihood ratio (LR) summarizes the statement of the forensic expert in the casework. However, the greatest interest to the jurists is the extent to which the *LR*s correctly discriminate "the same speaker and different-speaker" pairs under operating conditions similar to those of the case in hand. As was made clear in the US Supreme Court decision in Daubert case (Daubert v. Merrell Dow Pharmaceuticals, 1993) it should be criterial for the admissibility of scientific evidence to know to what extent the method can be, and has been, tested.

The principle for evaluation of the strength of evidence consists in the estimation and the comparison of the likelihood ratios that can be obtained from the evidence E, on one hand when the hypothesis H_0 is true (the suspected speaker truly is the source of the questioned recording) and, on the other hand, when the hypothesis H_1 is true (the suspected speaker is truly not the source of the questioned recording) [14]. The performance of an automatic speaker recognition method is evaluated by repeating the experiment described in the previous sections, with several speakers being at the origin of the questioned recording, and by representing the results using experimental (histogram based) probability distribution plots such as probability density functions and cumulative distribution functions in the form of Tippett plots (Fig. 3a) [10, 14].

The way of representation of the results in the form of Tippett plots is the one proposed by Evett and Buckleton in the field of interpretation of the forensic DNA analysis [6]. The authors have named this representation "Tippett plot," referring to the concepts of "within-source comparison" and "between-sources comparison" defined by Tippett et al.

Forensic Speaker Recognition in Mismatched Conditions

Nowadays, state-of-the-art automatic speaker recognition systems show very good performance in discriminating between voices of speakers under controlled recording conditions. However, the conditions in which recordings are made in investigative activities (e.g., anonymous calls and wire-tapping) cannot be controlled and pose a challenge to automatic speaker recognition. Differences in the background noise, in the phone handset, in the transmission channel, and in the recording devices can introduce variability over and above that of the voices in the recordings. The main unresolved problem in FASR today is that of handling mismatch in recording conditions, also including mismatch in languages, linguistic content, and non-contemporary speech samples. Mismatch in recording conditions has to be considered in the estimation of the likelihood ratio [11–13]. Next step can be combination of the strength of evidence using aural-perceptive and acoustic-phonetic approaches (aural-instrumental) of trained phoneticians with that of the likelihood ratio returned by the automatic



Voice, Forensic Evidence of. Figure 3 (**a**) Estimated probability density functions of likelihood ratios; (**b**) Tippett plots corresponding to (**a**). © IEEE.

system [4]. In order for FASR to be acceptable for presentation in the courts, the methods and techniques have to be researched, tested and evaluated for error, as well as be generally accepted in the scientific community. The methods proposed should be analyzed in the light of the admissibility of scientific evidence (e.g., Daubert ruling, USA, 1993) [11].

Summary

The essay discussed some important aspects of forensic speaker recognition, focusing on the necessary statistical-probabilistic framework for both quantifying and interpreting recorded voice as scientific evidence. Methodological guidelines for the calculation of the evidence, its strength and the evaluation of this strength under operating conditions of the casework were presented. As an example, an automatic method using the Gaussian mixture models (GMMs) and the Bayesian interpretation (BI) framework were implemented for the forensic speaker recognition task. The BI method represents neither speaker verification nor speaker identification. These two recognition techniques cannot be used for the task, since categorical, absolute and deterministic conclusions about the identity of source of evidential traces are logically untenable because of the inductive nature of the process of the inference of identity. This method, using a likelihood ratio to indicate the strength of the evidence of the questioned recording, measures how this recording of voice scores for the suspected speaker model, compared to relevant non-suspect speaker models. It became obvious that particular effort is needed in the trans-disciplinary domain of adaptation of the state-of-the-art speech recognition techniques to real-world environmental conditions for forensic speaker recognition. The future methods to be developed should combine the advantages of automatic signal processing and pattern recognition objectivity with the methodological transparency solicited in forensic investigations.

Related Entries

- ► Forensic Biometrics
- ► Forensic Evidence
- ► Speaker Recognition, An Overview

References

1. Rose, P.: Forensic Speaker Identification. Taylor & Francis, London (2002)
2. Dessimoz, D., Champod, C.: Linkages between biometrics and forensic science. In: Jain, A., Flynn, P., Ross, A. (eds.) Handbook of Biometrics, pp. 425–459. Springer, New York (2008)
3. Nolan, F.: Speaker identification evidence: its forms, limitations, and roles. In: Proceedings of the Conference "Law and Language: Prospect and Retrospect", Levi (Finnish Lapland), pp. 1–19 (2001)
4. Rose, P.: Technical forensic speaker recognition: Evaluation, types and testing of evidence. Comput. Speech Lang. **20**(2–3), 159–191 (2006)
5. Meuwly, D.: Voice analysis. In: Siegel, J., Knupfer, G., Saukko, P. (eds.) Encyclopedia of Forensic Sciences, pp. 1413–1421. Academic Press, London (2000)
6. Drygajlo, A.: Forensic automatic speaker recognition. IEEE Signal Process. Mag. **24**(2), 132–135 (2007)
7. Robertson, B., Vignaux, G.: Interpreting Evidence. Evaluating Forensic Science in the Courtroom. John Wiley & Sons, Chichester (1995)
8. Champod, C., Meuwly, D.: The inference of identity in forensic speaker identification." Speech Commun. **31**(2–3), 193–203 (2000)
9. Aitken, C., Taroni, F.: Statistics and the Evaluation of Evidence for Forensic Scientists. John Wiley & Sons, Chichester (2004)
10. Drygajlo, A., Meuwly, D., Alexander, A.: Statistical methods and Bayesian interpretation of evidence in forensic automatic speaker recognition. In: Proceedings of Eighth European Conference on Speech Communication and Technology (Eurospeech'03), pp. 689–692 Geneva, Switzerland, (2003)
11. Alexander, A.: Forensic automatic speaker recognition using Bayesian interpretation and statistical compensation for mismatched conditions. Ph.D. thesis, EPFL (2005)
12. Gonzalez-Rodriguez, J., Drygajlo, A., Ramos-Castro, D., Garcia-Gomar, M., Ortega-Garcia, J.: Robust estimation, interpretation and assessment of likelihood ratios in forensic speaker recognition. Comput. Speech Lang. **20**(2–3), 331–355 (2006)
13. Alexander, A., Dessimoz, D., Botti, F., Drygajlo, A.: Aural and automatic forensic speaker recognition in mismatched conditions. Int. J. Speech Lang. Law, **12**(2), 214–234 (2005)
14. Meuwly, D., Drygajlo, A.: Forensic speaker recognition based on a Bayesian framework and Gaussian mixture modelling (GMM). In: Proceedings 2001: A Speaker Odyssey, The Speaker Recognition Workshop, pp. 145–150 Crete, Greece, (2001)
15. Alexander, A., Drygajlo, A.: Scoring and direct methods for the interpretation of evidence in forensic speaker recognition. In: Proceedings of Eighth International Conference on Spoken Language Processing (ICSLP'04), pp. 2397–2400 Jeju, Korea, (2004)

Voiced Sounds

The voiced speech is generated by the modulation of the airstream of the lungs by periodic opening and closing of the vocal folds in the glottis or larynx. This is used, e.g., for vowels and nasal consonants.

► Speech Production

Voiceprint

Voiceprint is another name for spectrogram. This name is usually avoided because of its association with voiceprint recognition, which is a highly controversial method of forensic speaker recognition, which exclusively uses visual examination of spectrograms.

► Voice, Forensic Evidence of

Volunteer Crew

The volunteer crew for a biometric test is the individuals that participate in the evaluation of the biometric and from whom biometric samples are taken.

► Test Sample and Size

W

Walk-through

► Iris on the Move™

Walk-up

► Iris on the Move™

Watermarking, Biometric

A specific type of digital watermarking that includes biometric information either in the watermark, the host data, or both. The Biometric watermarking systems have the additional requirement of not degrading the performance of the biometric system(s) which they protect. This characteristic, sometimes referred to as performability, can entail any effect on matching performance, image quality, or computational efficiency. Systems watermarking biometric host data with biometric feature vectors add the potential for multimodal authentication.

► Iris Digital Watermarking

Watermarking, Digital

Digital watermarking is a method of embedding information within digital media for the purpose of proving file authenticity, tracking chain of custody and data reproduction, or describing host content. In digital watermarking, the watermark and the host data are typically related, and both are utilized by an intended recipient. Although not necessary, digital watermarks are typically imperceptible to humans either visually or audibly unlike their predecessors, paper watermarks. As a result, digital watermarking systems rely on machines to carry out the processes of watermark detection and extraction.

► Iris Digital Watermarking

Wavefront Coded® Iris Biometric Systems

V. Paúl Pauca[1], Kelly Smith Faddis[2], Arun Ross[2], Joseph van der Gracht[3], Todd C. Torgersen[1]
[1]Wake Forest University, Winston-Salem, NC, USA
[2]West Virginia University, Morgantown, WV, USA
[3]Holospex Inc., Columbia, MD, USA

Synonyms

Computational iris recognition systems; Pupil phase engineered iris biometrics

Definition

Wavefront coded iris biometrics is an imaging method whereby a suitably designed phase mask, placed in the pupil of an imaging system, is used to encode the depth dimension of an extended three-dimensional scene by means of an approximately shift-invariant point spread function. Iris images so acquired are thus deliberately distorted by a known amount in a way that is insensitive to misfocus blur, within a range

© 2009 Springer Science+Business Media, LLC

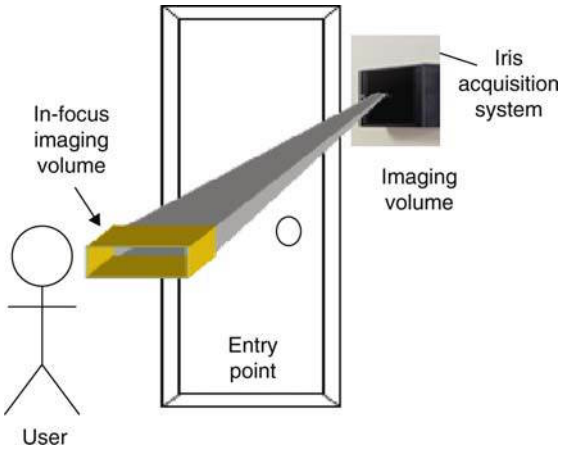
greater than the system's depth of field. For sufficient SNR, these images can be digitally deblurred by standard image deconvolution algorithms to recover depth dependent detail, enabling accurate iris recognition and identification. It is shown that even the intermediate, distorted iris images maintain sufficient low- and mid-frequency information to enable increased iris recognition performance, without resorting to digital restoration.

Introduction

The iris is a popular biometric that is gaining increased attention as a means of identification and verification of individuals for controlling access to secured areas, materials, or systems. The popularity of the iris as a biometric stems from its availability for remote and noninvasive assessment and the uniqueness of its texture from one subject to another, making possible a fully automated recognition and verification system based upon machine vision [1]. This texture is well known to provide a signature that is unique to each subject. In fact, the operating probability of false identification by the Daugman algorithm [2] can be of the order of 1 in 10^{10}. Compared with other biometric signatures, the iris is generally considered to be more stable and reliable for identification.

However, a major limitation of conventional iris recognition systems [3] is the inability to obtain in-focus iris images over an extended distance range. In order to achieve reasonable lighting levels and exposure times, the optical system must have a high numerical aperture and a corresponding low F-number, to provide sufficiently high signal-to-noise ratio (SNR) at the detector, with minimum motion blur. Unfortunately, a high numerical aperture results in a corresponding small depth-of-field and small ▶ iris recognition operating range. Often times, end-users are forced to "play the trombone" in order to present their iris within the system's imaging volume. Figure 1 illustrates a typical iris recognition scenario, where a user must submit a sample of his iris to the acquisition system, to gain access to a secured environment. The shaded box indicates the imaging volume that produces an in-focus acceptable image for processing.

Wavefront coding is a novel technology that was recently proposed for facilitating interaction of end-users with an iris biometric system [4, 5]. In

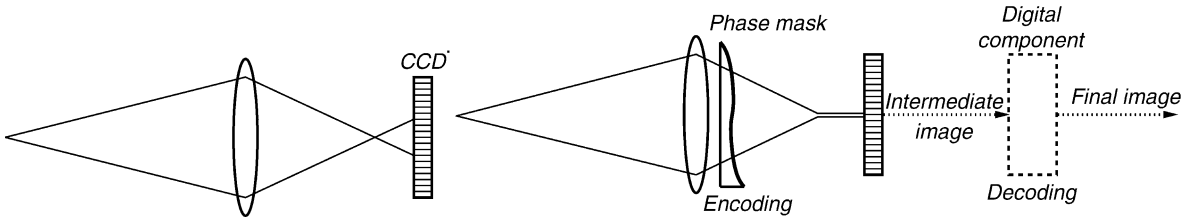


Wavefront Coded® Iris Biometric Systems. Figure 1 Iris recognition imaging volume.

particular, wavefront coding enables the acquisition of iris data through a large field of view and large depth of field using relatively fast optical systems. They do so without compromising optical resolution and light gathering capacity; a performance that cannot be achieved using traditional optical designs [4]. Wavefront coding was originally proposed by Dowski and Cathey [6], fitting within the concept of modern computational or task-based imaging. Since its original proposal, wavefront coding has been extended to include more general separable and non-separable type surfaces or phase masks [7, 8]. A more recent development includes a novel general framework, known as ▶ pupil phase engineering (PPE), to address high quality image acquisition from a numerical optimization perspective [9, 10]. In this framework, image quality requirements may include extending the depth of field, controlling or minimizing the impact of aberrations, motion blur, scattering from the imaging medium, among others (see, e.g., [10] and references therein).

Wavefront Coding for Extended Focus and Aberration Correction

Wavefront coding is a novel imaging modality where a unique aperture configuration is used to increase the depth of field, without significantly decreasing the SNR and light gathering capacity. In this modality, a phase mask is placed in the pupil of an imaging



Wavefront Coded® Iris Biometric Systems. Figure 2 Standard (*left*) and cubic phase (*right*) imaging systems.

system to deliberately distort the image, but in a way that is relatively insensitive to misfocus. Figure 2 illustrates the difference between conventional limited-focus and wavefront encoded systems. The distorted image carries depth-dependent intensity information which can be digitally recovered using standard deconvolution methods. The most general polynomial form for the added pupil phase mask $\phi(x,y)$ is the following:

$$\phi(x,y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}x^m y^n, \qquad (1)$$

The original proposal [5] argued for a cubic phase mask $\phi(x,y) = \alpha(x^3+y^3)$, based on the requirement that the encoding pupil phase $\phi(x,y)$ be monomial in the pupil coordinates (x,y), be separable in those coordinates, and lead to an MTF that is asymptotically focus-independent. Several other separable and non-separable, symmetric, and odd phase masks of the form (1) have been proposed and numerically optimized [11, 12] to achieve desirable characteristics in the point spread function (PSF). These may include insensitivity to defocus as well as the control of optical system aberrations such as spherical aberration, astigmatism, and curvature (see [10] for details regarding the numerical optimization of such phase masks using information theoretic based metrics).

The use of phase masks in iris recognition imaging systems is a promising approach that could greatly extend the *operational range* of iris recognition [4, 5], thereby facilitating flexibility when an end-user interacts with the system. The performance of wavefront coded iris biometric systems has been recently evaluated using small iris datasets (less than 10 user irises). A more comprehensive study using simulated unrestored wavefront coded images evaluated 150 iris images pertaining to 50 subjects has been recently conducted [13].

Simulation of Iris Biometric Systems

Fourier optics provides a convenient first order approximation of image formation and computer simulation systems often take advantage of this theory for efficient design and study of optical systems. The Simulator of Iris Recognition Imaging Systems (SIRIS) is one such a tool that was designed to generate conventional blurred and wavefront coded imagery, for a wide variety of polynomial form phase masks. SIRIS employs Fourier optics based image formation, PSF, and noise models to leverage the exploration of separable and non-separable phase masks on iris recognition performance [14].

SIRIS uses an implementation of Daugman's algorithm and, among other functionalities, allows the user to specify optical characteristics and parameters of an imaging device. These parameters include focal length, object distance, pupil diameter, pixel pitch, and noise. Using the specified optical and detector parameters and input imagery, SIRIS generates corresponding output imagery that includes the effects of a pre-specified phase mask and noise characteristics. It also provides the waves of defocus blur corresponding to any specific distance from the plane of best focus. The result is a model that summarizes the effects of defocus blur and the phase mask on the imaging system.

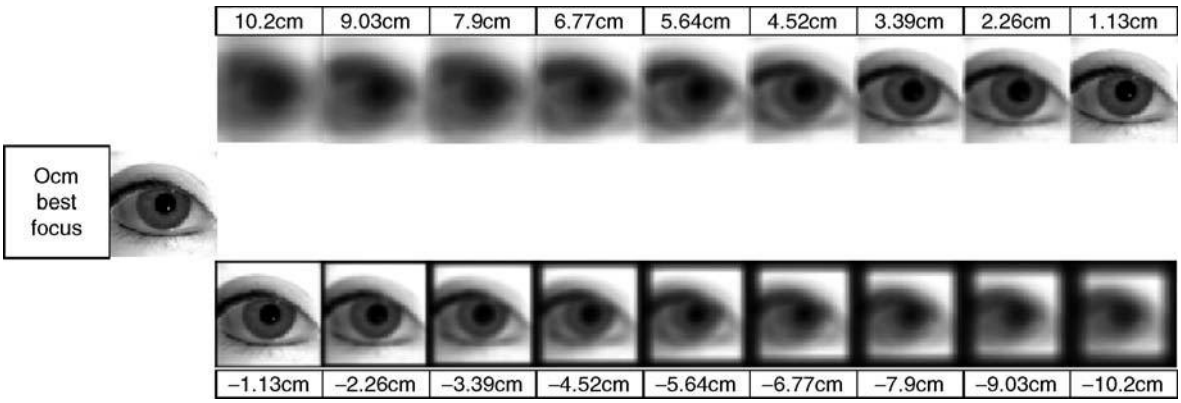
Simulation and Study of Unrestored Wavefront Coded Iris Imagery

SIRIS was recently employed to simulate, study, and compare the performance of conventional and wavefront coded iris biometric systems. A set of 150 in-focus iris images from the Iris Challenge Evaluation (ICE) dataset [15] were used in this study, pertaining to 50 different hand-selected users. The selected 150

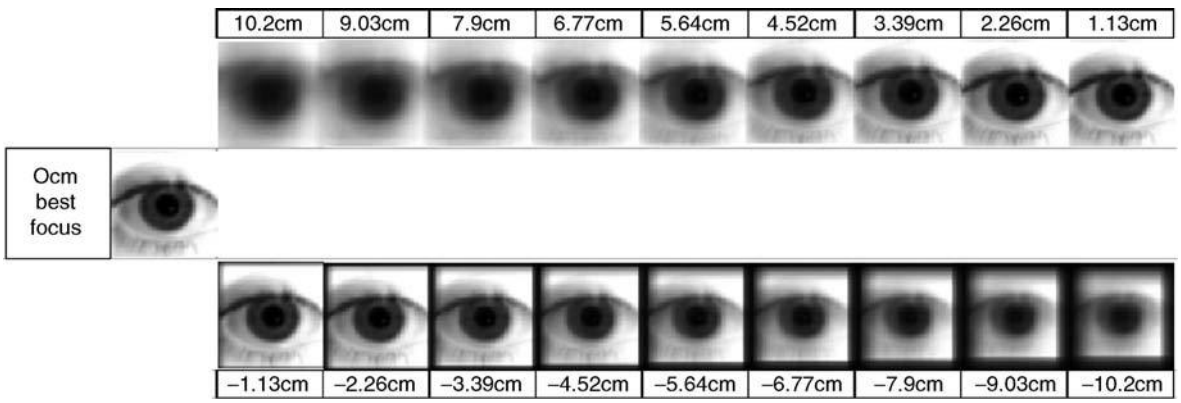
in-focus iris images were used as ground truth (reference images) in the iris recognition process, as well as an input to SIRIS for the generation of all conventional out-of-focus blurred and wavefront coded simulated imagery. Optical and detector parameters were selected based on the characteristics of the system used to collect the ICE images at Notre Dame University [15]. The F-number was set approximately to 1.92 for iris images placed at 0.5 *m* away from the detector. The pixel pitch is 5.134 μm. The wavefront coded system was modeled, having a cubic phase element with $\alpha = 30$. The system noise, having a Poisson signal dependent term associated with the light detection process and a white Gaussian independent term (1% noise), reflecting a camera under ideal lighting and image capture conditions was modeled.

The blurred images for both conventional and wavefront coded systems were obtained by modeling the movement of an iris in 1.13 *cm* increments away and towards the imaging device, in a range from −10.2 *cm* (away) to 10.2 *cm* (towards) from the plane of best focus. This range is significantly wider than the known imaging volume of the standard system. Figure 3 shows a sample set of blurred images obtained with a conventional system for a single input iris image.

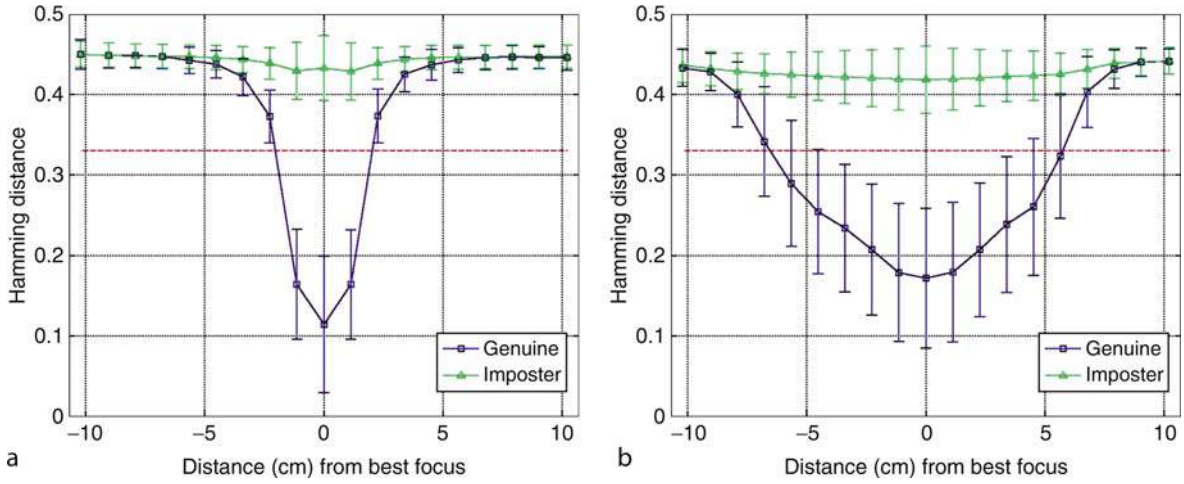
Notice that appropriate magnification of the iris due to movement away and towards the camera is taken into account. Figure 4 shows a sample set of blurred images obtained with a wavefront coded system ($\alpha = 30$) for a single input iris image. The wavefront coded images are unrestored, that is, no digital filter was applied to reverse the effects of the cubic phase mask. A total of 427,500 simulated conventional and wavefront coded images were generated from the 150 input iris images.



Wavefront Coded® Iris Biometric Systems. **Figure 3** Sample SIRIS simulated standard images.



Wavefront Coded® Iris Biometric Systems. **Figure 4** Sample SIRIS unrestored cubic images.



Wavefront Coded® Iris Biometric Systems. Figure 5 Imposter and genuine score statistics of standard (above) and unrestored cubic imagery (below) in which the enrollment images are at the plane of best focus.

Wavefront Coded® Iris Biometric Systems. Table 1 Operational range comparison of a conventional versus a wavefront coded system

System type	Distance (cm)	Operational range
Conventional	−2–2.5	∼4.5
Wavefront coded	−7–5.8	∼12.8

The blurred images were then processed using a research implementation of Daugmans algorithm. Specifically, the Hamming distances for all blurred images with respect to all reference images were computed. The enrollment images do not contain defocus blur. For our dataset, the calculation of Hamming distances implies the computation of a total of $2 \times 150 \times 150 \times 19$ matching scores (≈855,000 comparisons). Figure 5 shows the results of the comprehensive analysis (users 1–50) of imposter and genuine Hamming distance scores for both the conventional and wavefront coded systems. The increased operational range for iris recognition evident from these plots is reported in Table 1.

These results are in accordance with the published literature [5, 7] and highlight the efficacy of the simulation tool in generating imagery useful for research in this field. Future work will introduce image restoration of wavefront coded images in the iris recognition process. We expect restored images to produce smaller Hamming distances and further to increase the operational range of iris recognition of the wavefront coded imaging system, at the cost of increasing its computational requirements. However, the extension in the operational range of iris biometric systems using blurred wavefront coded imagery is an appealing prospect, since it does not increase the computational requirements of the system and processing time.

Summary

Depth of field and operational range play a major limiting role in the application of iris recognition technology. The addition of a wavefront coding phase masks was shown, via simulation, to significantly increase the operational range of iris recognition, even in the absence of restoration schemes. The tradeoff between degradation in Hamming distance scores and the inclusion of the cubic phase mask was also demonstrated. The utility of the simulation software has been anecdotally verified against actual blurred and wavefront coded imagery, an investigation on the use of wavefront coding imaging technology on a large image collection is needed. Future initiatives would consider the development of novel pattern recognition algorithms to capitalize on the information provided by a wavefront coded element. The use of segmentation methods that are less reliant on sharp pupil/iris and iris/sclera boundaries will also help to improve the performance. This paper has shown that wavefront coding technology can improve the operational range associated with iris recognition systems accompanied by a

reduction in SNR that does not dramatically affect the performance of the system. Additional work in systems engineering and performance optimization is necessary in order to reap the benefits of this technology.

Related Entries

- ▶ Biometric Sample Acquisition
- ▶ Biometric System Design
- ▶ Iris Encoding and Recognition

References

1. Daugman, J.G.: High confidence visual recognition of person by a test of statistical independence. IEEE Trans. PAMI **15**, 1148–1161 (1993)
2. Daugman, J.: How iris recognition works. IEEE Trans. Circuits Syst. Video Technol. **14**(1), 21–30 (2004)
3. Bowyer, K., Hollingsworth, K., Flynn, P.: Image understanding for iris biometrics: A survey. Tech. rep., University of Notre Dame, CSE (2007)
4. Narayanswamy, R., Silveira, P.E.X., Setty, H., Pauca, V.P., van der Gracht, J.: Extended depth-of-field iris recognition system for a workstation environment. In: Proceedings of SPIE Biometric Technology for Human Identification II, vol. 5779. Orlando, FL (2005)
5. van der Gracht, J., Pauca, V.P., Setty, H., Jr., E.R.D., Plemmons, R.J., Torgersen, T.C., Prasad, S.: Iris recognition with enhanced depth-of-field image acquisition. In: Proceedings of SPIE Defense and Security Symposium, vol. 5358. SPIE, Orlando, FL (2004)
6. Dowski, E.R., Cathey, W.T.: Wavefront coding for detection and estimation with a single-lens incoherent optical system. In: Proceedings of Acoustics, Speech, and Signal Processing, vol. 4, pp. 2451–2454 (1995)
7. Dowski, E.R., Cormack, R.H., Sarama, S.D.: Wavefront coding: Jointly optimized optical and digital imaging systems. In: Proceedings of AeroSense Conference. Orlando, FL (2000)
8. Cathey, W.T., Dowski, E.R.: New paradigm for imaging systems. Appl. Opt. **41**(29), 6080–6092 (2002)
9. Prasad, S., Torgersen, T.C., Pauca, V.P., Plemmons, R.J., van der Gracht, J.: High-resolution imaging using integrated optical systems. Int. J. Imaging Syst. Technol. **14**(2), 67–74 (2004). Special Issue: High-Resolution Image Reconstruction I. (invited paper)
10. Prasad, S., Pauca, V., Plemmons, R., Torgersen, T., van der Gracht, J.: Pupil-phase optimization for extended-focus aberration-corrected imaging systems. In: Proceedings of SPIE, Advanced Signal Processing Algorithms, Architectures, and Implementations XIV, vol. 5559, pp. 335–345. SPIE, Denver, CO (2004)
11. Prasad, S., Torgersen, T., Pauca, V., Plemmons, R., van der Gracht, J.: Engineering the pupil phase to improve image quality. In: Z. Rahman, R. Schowengerdt, S. Reichenbach (eds.) Proceedings of SPIE Visual Information Processing XII, vol. 5108, pp. 1–12 (2003)
12. Pauca, V., Plemmons, R., Prasad, S., Torgersen, T., van der Gracht, J.: Integrated optical-digital approaches for enhancing image restoration and focus invariance. In: Proceedings of SPIE, Advanced Signal Processing Algorithms, Architectures, and Implementations XIII, vol. 5205, pp. 348–357. SPIE, San Diego, CA (2003)
13. Smith, K.N., Pauca, V.P., Ross, A., Torgersen, T.C., King, M.C.: Extended evaluation of simulated wavefront coding technology in iris recognition. In: Proceedings of IEEE Conference on Biometrics: Theory, Applications and Systems. SPIE, Washington, DC (2007)
14. Leonhardt, E., Pauca, V.P.: A User Guide for SIRIS: Simulator of Iris Recognition Imaging Systems. Wake Forest University, CS (2004)
15. Phillips, P.J., Scruggs, W.T., OToole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M.: Frvt 2006 and ice 2006 large-scale results. Tech. Rep. NISTIR 7408, National Institute of Standards and Technology (2007)

Wavefront Coding

Wavefront Coding is a computational imaging technique that relies on the simultaneous optimization of the pupil function and the image processing of a digital imaging system to increase its tolerance to optical aberrations. When the aberration being corrected is defocus, Wavefront Coding is able to provide the system designer with an extension of the depth of field of the imaging system.

▶ Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition

Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition

Paulo E. X. Silveira, Lu Gao,
Ramkumar Narayanswamy
Omni Vision CDM Optics, Boulder, CO, USA

Synonym

Wavefront Coded® (Wavefront Coded® is a registered trademark of Omni Vision CDM Optics, Inc.) biometric iris recognition

Definition

Wavefront Coding™ (Wavefront Coding™ is a trademark of Omni Vision CDM Optics, Inc.) is a computational imaging technique capable of increasing the depth of field of an imaging system and increasing the amount of information captured within a given imaging volume. Iris recognition consists of using the unique texture present in the human iris to perform biometric identification. The small size of the texture details compared to the subject standoff distance and normal human motion makes it challenging to capture high quality iris images using traditional imaging techniques. The use of Wavefront Coding alleviates these requirements and shows great performance improvements by reducing the constraints in subject position while maintaining short exposure times during image capture.

Introduction

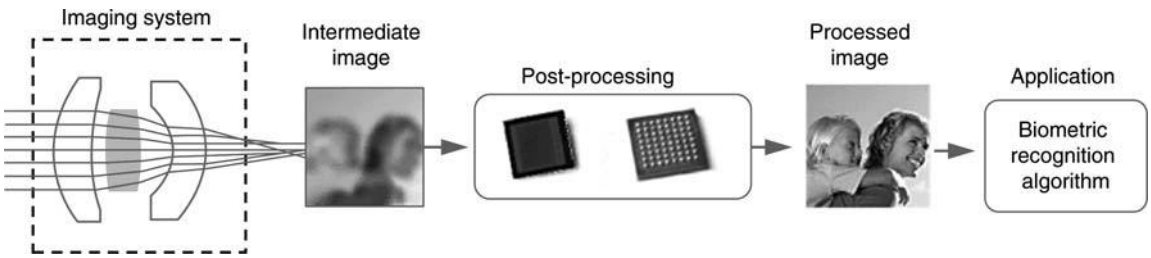
The human iris has several unique features that make it an attractive choice for biometric recognition: (1) the iris texture is highly unique (possesses high entropy); (2) the texture is stable over the lifetime of an individual; and (3) the image of the eye and iris can be captured from a distance using standard imaging systems [1–3]. A wide range of imaging systems including hand-held cameras, miniature cameras, telephoto, and multi-spectral cameras are currently being considered for use or are used to capture the image data for iris recognition. These imaging systems are characterized by an ► imaging volume over which the image of the iris is captured with adequately high signal quality. The challenge faced by today's iris recognition systems is that subjects must find this rather small imaging volume by trial-and-error. Then, the iris to be recognized must remain stationary during the image capture time to avoid motion blur. Thus, the overall subject experience is not satisfactory and needs improvement.

Ideally, iris capture systems should operate well with stationary subjects as with moving subjects, and the image quality should remain high over the entire imaging volume. This requirement calls for a relatively fast optical system (low F/#) which, in traditional imaging systems, inherently leads to a shallow depth of field. On the other hand, an ideal system should have a large field of view and a large depth of field so that subjects can be identified with minimal cooperation [4, 5]. These seemingly conflicting requirements can be satisfied by using ► Wavefront Coding, a computational imaging technique that is capable of increasing the depth of field of an imaging system without needing to increase its F-number. Note that the traditional method of stopping down the aperture to increase the F-number is detrimental to iris recognition in two ways: (1) the brightness of the image, which is proportional to the signal level, drops proportionally to the area of the obscured aperture; and (2) the optical resolution drops proportionally to the reduction in diameter of the aperture.

Wavefront Coded Iris Recognition System

Wavefront Coding relies on the simultaneous optimization of the pupil function and the image processing of a digital imaging system to increase its tolerance to aberrations that would typically degrade the image [6]. For example, when the aberration being corrected is defocus, Wavefront Coding is able to provide the system designer with an extension of the depth of field of the imaging system [7]. This is done by using non-conventional aspherical optics to produce point spread functions (PSFs) that do not vary significantly over an extended imaging volume compared to traditional optics. Note that the pupil function can be optimized by inserting a new element at the exit pupil or, more practically, by altering the shape of existing optical surfaces. When the goal is to increase the imaging volume used for biometric iris recognition, Wavefront Coding can be employed to provide a modulation transfer function (MTF) of the system that is consistent over the required range of spatial frequencies and over a wide range of field points and defocus [8].

Figure 1 shows the main functional blocks of a Wavefront Coded biometric recognition system. Typically, the imaging system consists of a low-F/# digital imaging system in which the pupil function is optimized to produce PSFs that are mostly invariant with defocus and MTFs that retain the required information over a sufficiently large range of spatial frequencies. A digital sensor captures an intermediate (Wavefront Coded) image and a processed image is generated by post-processing this image. The processing may be as simple as a linear filtering step, with filters that are



Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition. **Figure 1** Block diagram of a Wavefront Coded biometric recognition system. The Wavefront Coded imaging system captures an intermediate image, which is decoded during post-processing. The processed image is then passed along to the biometric recognition algorithm.

designed in tandem with the Wavefront Coded optics and have as a goal the conversion of the defocus-invariant PSFs into diffraction-limited PSFs. The filtering step consists of a small fraction of the processing steps required for conventional iris recognition algorithms and, therefore, contributes with minimal processing overhead (e.g., see [1–3] for comparison).

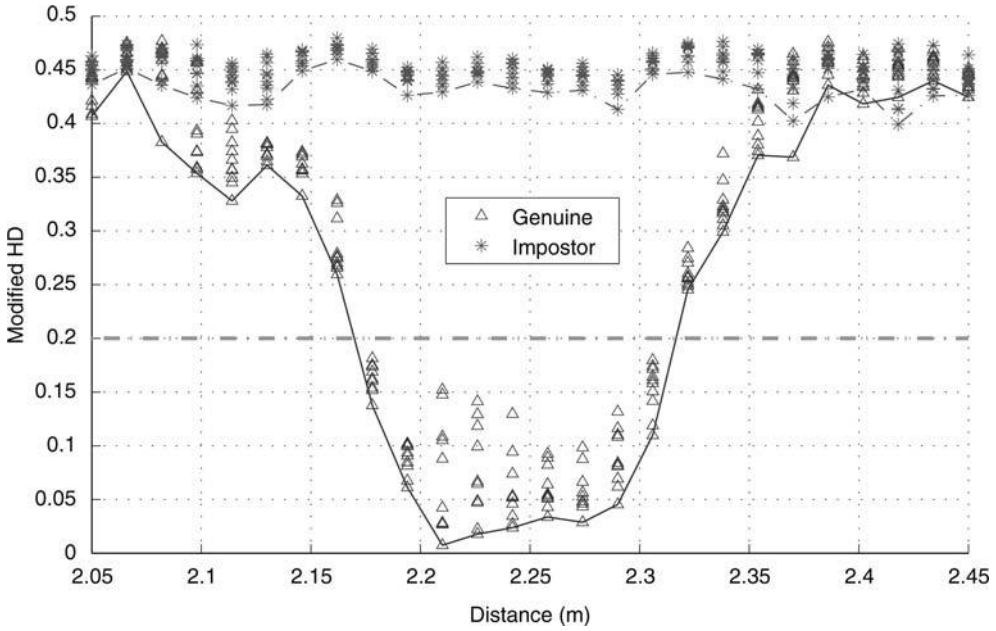
The resulting processed image has an extended depth of field and is provided as an input to a biometric recognition algorithm that uses the same range of spatial frequencies used in the design of the modified imaging system including the Wavefront Coded element and deconvolution filter. An example of experimental results obtained using a Wavefront Coded biometric recognition system for performing robust iris recognition is detailed below.

Experimental Example

In this example, the performance of a traditional (non-Wavefront Coded) iris recognition imaging system is compared with that obtained with a Wavefront Coded imaging system. In both cases the subject stands at a distance of about 2.2 m from the imaging system and the performance is compared in terms of the range of subject distances over which iris recognition can be correctly performed. In both cases the image of the subject's iris is captured and compared with that of a pre-enrolled iris, providing a match in terms of a modified (normalized) Hamming distance (HD). The HD is a measure of the number of bits that are different between any two given ► iris codes. A HD of zero represents a perfect match while a HD of 0.5 is equivalent to randomly varying bits and, therefore, corresponds to a complete mismatch. In practical applications, a HD below 0.2 or 0.3 corresponds to a match between the image of a captured iris and a pre-enrolled image.

Figure 2 shows a plot of the HD as a function of the subject position. At each position ten images of two eyes (a genuine eye and an impostor eye) are captured in rapid succession. The images are then processed and compared to a pre-enrolled iris code. All the calculated HDs for the genuine (impostor) iris are plotted as triangles (stars) and the minimum HDs are connected by a solid (dotted) line. Ten images were collected in order to provide us with an estimate of the statistical variation of the HD at each position. In reality, a single match (a HD below 0.2) would be sufficient to identify a subject as genuine, making it unnecessary to capture additional images. Because imperfect image segmentation is the main cause of variation in the HD, the minimum HDs are selected to aid in isolating this effect when comparing the performance of the two systems. Note that in the traditional system the solid line has the shape of a narrow valley with sharp transition regions, corresponding to the distances at which defocus causes a drop in modulation at the spatial frequencies used by the recognition algorithm. The flat region close to the best-focus position corresponds to the distances at which defocus does not impact the spatial frequencies of interest significantly. Beyond this range the image of the iris becomes so defocused that it is no longer possible to resolve the details necessary to correctly perform iris recognition. Using a threshold HD of 0.2, an operating range of about 16.5 cm is measured for this traditional imaging system.

Figure 3 shows a plot of the HD as a function of the iris distance for a Wavefront Coded system with



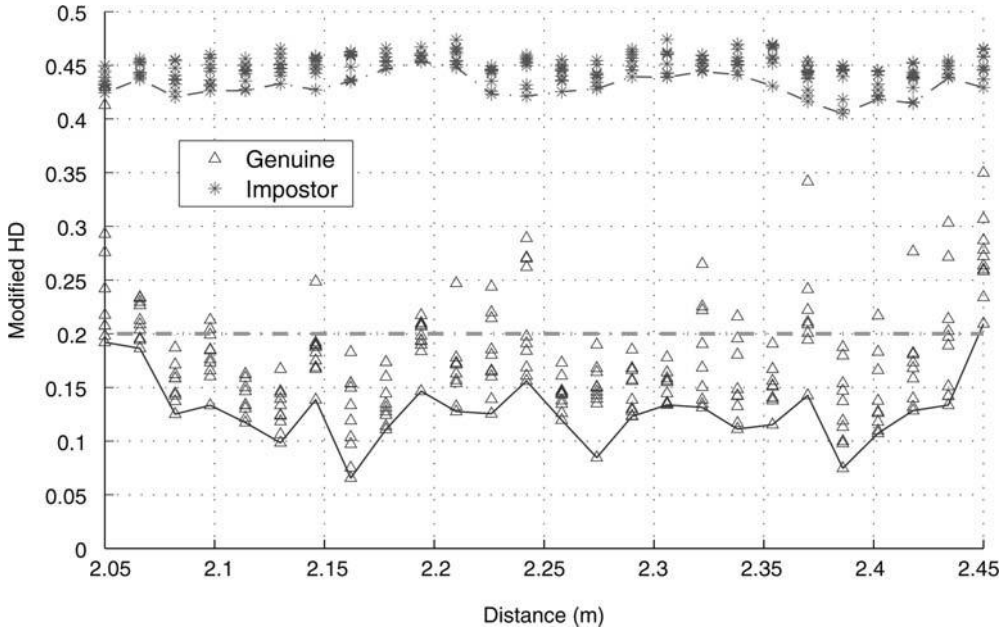
Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition. **Figure 2** Iris recognition range of a traditional imaging system. The iris recognition (measured in terms of a modified HD) is plotted as a function of subject distance showing a narrow recognition range (16.5 cm). Ten measurements are taken at each position and images of a genuine iris (triangles) are compared to those of an impostor iris (stars). A solid (dotted) line connects the lowest HD of the genuine (impostor) iris.

the same characteristics as that of the traditional imaging system providing the data for Fig. 2. The imaging system providing the data for Fig. 3 uses the same detector, same biometric algorithm, and same optical elements as the one used for the data of Fig. 2, except for the inclusion of a Wavefront Coded element at the aperture stop and a decoding step of the intermediate image before biometric processing. In this case, the solid line defines a shallow and broad valley, effectively demonstrating the trade-off of the lowest HD for an extended depth of field. In this example, Wavefront Coding provides us with a recognition range of almost 40 cm at a threshold HD of 0.2, providing an extension of the depth of field 2.4 times greater than that which was achieved with the traditional system while maintaining a sufficiently large discriminatory capacity between genuine and impostor irises. This extension in the depth of field can be tuned by the system designer and it translates into greater usability of the iris recognition system, possibly representing the difference between a useful system and a system that is so constrained that has little practical use. In this example, the Wavefront Coded element was fabricated separately to simplify the comparison of the traditional test system with the Wavefront Coded one. In practice, however, it would be more advantageous to simply modify the shape of one or more of the existing optical surfaces.

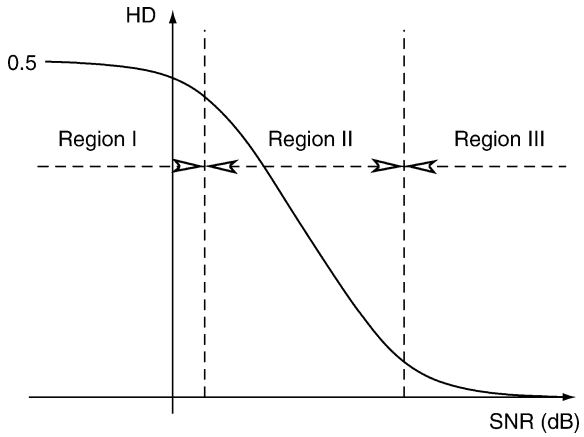
System Design

As mentioned above, the system designer can tune the extension of the depth of field to match the needs of the intended application. Therefore a discussion of the tradeoffs is warranted. The tradeoffs may be well understood in terms of the signal-to-noise ratio (SNR) available to the designer, as explained by the HD curve of a hypothetical biometric recognition system shown in Fig. 4.

When designing a Wavefront Coded imaging system, the system designer is given the choice of how much to reduce the SNR at best focus in order to convert it into an extended depth of field. As it turns out, this SNR can be precisely calculated [9] and it is directly proportional to the square of the MTF of the imaging system. Figure 4 shows a plot of the HD as a



Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition. **Figure 3** Iris recognition range of a Wavefront Coded imaging system. The iris recognition (measured in terms of a modified HD) is plotted as a function of subject distance showing a broad recognition range (40 cm). Ten measurements are taken at each position and images of a genuine iris (triangles) are compared to those of an impostor iris (stars). A solid (dotted) line connects the lowest HD of the genuine (impostor) iris.



Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition. **Figure 4** HD as a function of SNR for a hypothetical biometric recognition system. Note the existence of three regions of operation: (I) a low SNR region, where the SNR is so low that identification is not possible, (II) a mid-SNR region, where a decrease in the SNR is accompanied by a proportional decrease in the HD; III) a high SNR region, where a decrease in SNR produces a negligible decrease in the HD.

function of the SNR. In general, the HD is a non-linear but monotonically decreasing function of the SNR. The binomial distribution of error bits in the iris provides us with three distinctive regions of interest. In Region I, the low-SNR region, the SNR is sufficiently low that the HD is approximately constant at 0.5. In Region II, the mid-SNR region, the decrease in the SNR is proportional to the decrease in the HD. In Region III, the high-SNR region, the SNR is sufficiently high that a decrease in the SNR results in a negligible decrease in the HD. The existence of Region III is generally possible due to the non-linear processing used in the generation of the iris codes (e.g., the thresholding step that is typically used to produce the binary iris code). Region III is clearly the preferred region of operation, and it demonstrates the fact that it is possible to extend the depth of field of an imaging system without any noticeable degradation of the HD at best focus. Then, the job of the system designer consists of (1) selecting an initial imaging system that operates preferably in Region III; and (2) simultaneously optimizing the pupil function and the deconvolution filter so that the HD is maintained above a threshold.

It should be noted that this threshold is application-specific, and that the optimization should be simultaneously performed over the entire imaging volume.

Summary

Wavefront Coding can be used to increase the depth of field of an imaging system by making the system more tolerant to defocus. This may be done without reducing the light collection ability of the system or its optical resolution, as would be the case when using the traditional method of stopping down the aperture. Moreover, the additional steps required in image processing are a small fraction of the total number of steps required for biometric recognition, translating in an increase in system robustness with little overhead. The experimental example described above demonstrates how Wavefront Coding can be used to increase the depth of field of an iris recognition system and, by doing so, reduce the constraints on the subject's position and effectively increase the usability of the system. Finally, the tradeoff between the SNR of an imaging system and the extension of its depth of field has been presented and discussed in association with designing a Wavefront Coded system for a given biometric application.

Related Entries

- ▶ Biometric Sample Acquisition
- ▶ Biometric System Design, Overview
- ▶ Iris on the Move
- ▶ Iris Recognition, Overview

References

1. Daugman, J.G.: High confidence visual recognition of person by a test of statistical independence. IEEE Trans. PAMI **15**, 1148–1161 (1993)
2. Daugman, J.G.: The importance of being random: statistical principles of iris recognition. Patt. Rec. **36**, 279–291 (2003)
3. Daugman, J.G.: How iris recognition works. IEEE Trans. Circuits and Systems for Video Tech. **14**(1), 21–30 (2004)
4. Gracht, J. van der, Pauca, V.P., Setty, H., Dowski, E.R., Plemmons, R.J., Torgersen, T.C., Prasad, S.: Iris recognition with enhanced depth-of-field image acquisition. In: Proceedings of the SPIE 5358, pp. 120–129. Orlando, FL (2004)
5. Plemons, R., Horvath, M., Leonhardt, E., Pauca, P., Prasad, S., Robinson, S., Setty, H., Togersen, T., Gracht, J. van der, Dowski, E., Narayanswamy, R., Silveira, P.E.X.: Computational Imaging Systems for Iris Recognition. In: Proceedings of SPIE 5559, pp. 346–357. Denver, CO (2004)
6. Cathey, W.T., Dowski, E.: A new paradigm for imaging systems. Appl. Opt. **41**, 6080–6092 (2002)
7. Dowski, E., Cathey, W.T.: Extended depth of field through wavefront coding. Appl. Opt. **34**, 1859–1866 (1995)
8. Narayanswamy, R., Baron, A.E., Chumachenco, V., Greengard, A.: Applications of Wavefront Coded Imaging. In: Proceedings of the SPIE 5299, 163–174. Orlando, FL (2004)
9. Silveira, P.E.X., Narayanswamy, R.: SNR Analysis of Task-based Imaging Systems with Defocus. Appl. Opt. **45**, 2924–2934 (2006)

Wavelength

Electromagnetic radiation includes visible light, ultraviolet light, infrared radiation, and radio waves. In classical electromagnetic field theory, electromagnetic radiation propagates as a pair of coupled electrical and magnetic fields. A time varying electrical field creates a time varying magnetic field that in turn creates a time varying electrical field that creates a time varying magnetic field that creates a time varying electrical field, and so on. The distance over which the electrical (or magnetic) field at a given instant varies from its maximum to its minimum and back is the wavelength of the radiation. Interactions between radiation and matter depend on wavelength. In the human visual system, wavelength is intimately connected with color. A wavelength of the order of 440 nm is perceived as blue, 740 nm as red and the human eye is relatively insensitive to wavelengths longer than 800 nm. On fundamental physical grounds, the resolution of optical systems is generally limited to features of the order of the wavelength of the light used. There are some special cases in which resolution can be extended beyond these limits.

- ▶ Biometirc Sample Acquisition
- ▶ Biometric Sensor and Device, Overview

Wavelet Transform

Wavelet transform divides a given data into different frequency components at different scales. In biometrics,

operating in wavelet domain provides additional edge and frequency information useful for fusion and recognition.

► Fusion, Sensor Level

Web Services

The W3C defines a Web service as "a software system designed to support interoperable Machine to Machine interaction over a network." Web services are frequently just Web APIs that can be accessed over a network, such as the Internet, and executed on a remote system hosting the requested services.

► Biometric Interfaces

Work-sharing On-card Matching

The reader side executes part of the calculation (such as alignment) to assist the smartcard to have a faster biometric verification time. The final matching score calculation shall be executed on-card.

► On-Card Matching

World Model

► Universal Background Models

X

XML Schema Definition

XML Schema definition is a notation (using the XML syntax) for defining the form of an XML-encoded message (similar to ASN.1, but with a different syntax and supporting only XML encoding of the data).

► Biometric Technical Interface, Standardization

© 2009 Springer Science+Business Media, LLC

Z

Zero Effort Forgery

Stephen J. Elliott
Purdue University, West Lafayette, IN, USA

Synonym

Zero effort impostor attempt

Definition

An impostor attempt is classed as "zero-effort" if the individual submits his/her own biometric feature as if he/she were attempting successful verification against his/her own template, but the comparison is made against the template of another user. In the case of dynamic signature verification, an impostor would therefore sign his/her own signature in a zero-effort attempt. In such cases where impostors may easily imitate aspects of the required biometric, a second impostor measure based on "active impostor attempts" may be required [1]. The active impostor attempt is the one in which an individual tries to match the stored template of a different individual by presenting a simulated or reproduced biometric sample or by intentionally modifying his/her own biometric characteristics [1].

Introduction

The concept of zero-effort forgery centers around an individual's capability to submit his/her own biometric feature as if he/she were attempting a successful verification against his/her own template. Zero-effort is a typical methodology of extracting performance rates (false accept, false reject) of a biometric system. However, in the realm of dynamic signature verification, the concept of the impostor distribution, hence zero effort, is often overlooked. This definition is broken into two parts:an overview of dynamic signature verification (to explain the concept of the modality that has the exception for zero effort) and a discussion of the importance of the impostor dataset (as it relates to dynamic signature verification and the concept of zero-effort as explained earlier).

Overview of Dynamic Signature Verification

Biometrics can be divided into two main categories: behavioral andphysiological. Fingerprint and iris are examples of physiological biometrics, whereas voice and signature are examples of behavioral biometrics. Signature verification aims at authenticating an individual, based on their signing characteristics. Applications such as document authenticity, financial transactions, and paper-based transactions have all, at one time, used the signature to convey the intent to complete a transaction [2,3]. Signature verification can be divided into two classes, depending on how the signature data are collected and analyzed. The first is the digitized signature. Here, the input is a static image, extracted from a document, check, or receipt that is verified after being scanned. The second type is a dynamic signature, where the user signs on a digitizer, and the signature, along with traits or characteristics, is collected in real time.

Dynamic signature verification has a number of statistical features that can be derived from the basic set of data that a digitizer provides, and these vary significantly across algorithms. The typical consensus is that these input variables are gathered from a digitizer and include x, y (Cartesian co-ordinates), p (pressure or force), and t (time) [4]. These input variables are then used to create the global and local features

© 2009 Springer Science+Business Media, LLC

described in various accounts in the literature [4–6]. Fairhurst and Ng [7] outlined 61 features [8], Nalwa [9], observed that the temporal "characteristics of the production of an online signature are the key to the signature's verification" (p. 5). Typical signature functions include pressure vs. time, horizontal and vertical components of position, velocity, acceleration, and force, all against time. Another way of characterizing a signature is through the analysis of the "stroke" that is, the pen down–pen up movement of the pen on the digitizer. All these various dynamic traits that are collected during the act of signing are said to make an impostor signature easier to detect.

Forgery Levels – The Impostor Data Set

One of the interesting characteristics of dynamic signature verification is the concept of an impostor attempt. An impostor attempt is defined as a "zero-effort" if the individual submits his/her own biometric feature as if he/she were attempting successful verification against his/her own template, but the comparison is made against the template of another user [1]. This makes sense for an opportunistic forger for almost all biometrics, barring dynamic signature verification. In the case of dynamic signature verification, "B" would sign "B"'s name trying to get into "A"'s account. Thus, zero-effort forgery, on the part of the imposter, makes no sense. "B" would not have any information about the signature that "A" provides. Therefore, in order to examine the FAR and FRR of a dynamic signature verification system, an examination of the impostor dataset construction is of great importance. In the literature, there are several different examples of the impostor dataset. Biometric testing and evaluation typically create a probability of an impostor signature getting accepted as a genuine signature. These outcomes are generally denoted by a False Accept, or a False Reject under conditions of zero effort. Such an attempt is defined as "where an impostor uses his or her own biometric sample and claims the identity of a different enrollee" [10].

The determination of the forger in dynamic signature verification is somewhat a challenge, as a zero effort forgery would require the impostor to sign his or her own name while claiming the identity of a different enrollee. Therefore, understanding how to test and evaluate the signature in a forgery setting is an interesting research question. In the literature, there is no real consensus on how best to challenge a signature. Impostor datasets are created in numerous ways and have the effect of changing the respective performances of algorithms. This change can be done through a different generation of the impostor signatures. A review of the literature shows various performance results from several studies, all of which have different methodologies for collecting impostor signature datasets. The various studies and their respective error rates (false accept, false reject, and the equal error rate where appropriate) have been outlined earlier [11]. The variances in error rates shown (0% to 50% false accept rate and 0% to 20% false reject rate) can be explained by a number of factors, one of which has to do with how an imposter signature dataset is created. Komiya and Matsumoto's [12] study is particularly interesting. This study had a database consisting of 293 genuine signatures and 540 forgery signatures from eight individuals [13] study dataset was comprised of 496 original signatures from 27 people. Each person signed 11 to 20 times. The database contained 48 forgeries that "fulfill the requirement on the visual agreement and the dynamic similarity with the original signature" (p. 5). [14] trained the algorithm using 250 signatures per writer; of these 250 signatures, 100 were authentic signatures and 150 were random forgeries, classified as the genuine signatures of other writers. [15] used 27 people in their study, with the participants writing their own signature. The study also used 4 people who imitated the signatures of these 27 people. Unfortunately, no further information is provided on the selection of the impostor or on what knowledge the forgers possessed in order to forge the signatures. [2] used genuine signatures from other individuals as forgeries. In addition, a group of synthesized signatures was created by distorting real signatures through the addition of low level noise and dilation or erosion of the various structures of the signature. [16] motivated the forgers by offering a cash reward. [17] examined people's signatures over a four-month period to assess variability over time. In the Signature Verification Competition, genuine signers created signatures other than their own [18]. In [19], the authors defined three different levels of forgeries: the simple, statically skilled, and

timed (p. 643). [3] used signatures that "on casual visual inspection would pass as authentic" (p. 201). [20] provides three characteristics of forgery: the random forgery defined as one that belongs to a different writer of the signature model; simple forgery represented by a similar shape consistency with the genuine signers shape; and the skilled forgery (p. 2). [8] described two classes, the skilled forgery and the random or zero effort forgery. For the skilled forger, it is "signed by a person who has access to a genuine signature for practice". In addition, it is important to understand the handedness of the genuine signer, and sometimes, the gender. Again, the literature sometimes provides insights into the impostor, and to some extent, the genuine datasets, all to varying degrees. For example, Lee, Berger, Aviczer [19] collected 5,603 genuine signatures from 105 people, including 22 women and 5 individuals who were left handed.

Therefore, when assessing the impostor datasets for dynamic signature verification, there are probably many different levels from "zero-effort" at the lowest level through to trained and skilled forger. This all depends on the level of information given to the impostor. For example, level 0 would be the "zero-effort" where "B" has no relevant knowledge of "A". The knowledge is zero; there is no knowledge of name: verbal, printed, or signed. The next level is that "B" knows "A"'s name – this information is revealed verbally. Therefore, the forger might mistake "Stephen" for "Steven" or "Steve," depending on the formality or ceremony of the signature. Level 2 would be that "B" has seen a static image of "A"'s signature prior to the impostor attempt. Level 3 is that "B" can see a static image of "A"'s signature at the time of signing. Level 4 is that "B" is able to trace a sample of "A"'s signature. Level 5 is where "B" has recently witnessed a sample of "A"'s signing, and level 6 is where "B" has repeatedly witnessed "A" signing. Even this methodology takes into consideration neither the Perceived Signature Strength (PSS) of the original signature nor the motivation of the impostor. The PSS is a concept that indicates that an "opportunistic forger" will not forge a signature that is "hard" to forge, as their success at the point-of-sale may be not as high as an "easy" signature. This is more of an "opportunistic" forgery as opposed to a more sophisticated attack on the signature as outlined in previous research. For the case of this study, an opportunistic forger is analogous to an "opportunistic thief" – working on their own without any equipment [16].

References

1. Mansfield, A and National Physical Laboratory (Great Britain): Best practices in testing and reporting performance of biometric devices (2002)
2. Hamilton, D., et al.: Low cost dynamic signature verification system. European Convention on Security and Detection, pp. 202–206 (1995)
3. Nelson, W., Kishon, E., Use of dynamic features for signature verification. Decision Aiding for Complex Systems, Conference Proceedings, 1991 IEEE International Conference on Systems, Man and Cybernetics, vol.1, pp. 201–205 (1991)
4. Vielhauer, C., Steinmetz, R., Mayerhofer, A.: Biometric hash based on statistical features of online signatures. Proceedings of the 16th International Conference on Pattern Recognition, Quebec City, Canda, vol. 1, pp. 123–126 (2002)
5. Leclerc, F., Automatic signature verification: the state of the Art–1989–1993. Int. J. Pattern Recog. Artif. Intell. **8**, 643–660
6. Elliott S., Hunt, A.: Dynamic Signature Forgery Assessment and Signature Strength Perception, Aeros. Electron. Syst. Mag. **23**, 13–18 (2008)
7. Fairhust, M., Ng, S.: Management of access through biometric control: A case study based on automatic signature verification, Universal Access in the Information Society, vol. 1, June 2001, pp. 31–39
8. Kholmatov, A., Yanikoglu, B.: Biometric authentication using online signatures. Computer and Information Sciences (ISCIS) 2004, pp. 373–380; http://www.springerlink.com/content/a7q4keh1dv8d1u5f
9. Nalwa, V.: Automatic on-line signature verification, Computer Vision – ACCV'98, 1997, pp. 10–15; http://dx.doi.org/10.1007/3–540–63930–6_98
10. Ohishi, T.: et al.: Pen-input on-line signature verification with position, pressure, inclination trajectories, Parallel and Distributed Processing Symposium, Proceedings 15th International, 2001, pp. 1757–1763
11. Elliott, S., Hunt, A.: An assessement of dynamic signature forgery and perception of signature strength. Procedding of the 40th Annual IEEE International Carnahan Conference on Security Technolgy, pp. 186–190 (2006)
12. Komiya Y., Matsumoto, T.: On-line pen input signature verification PPI (pen-Position/pen-Pressure/pen-Inclination), IEEE SMC '99 Conference Proceedings. IEEE International Conference on Systems, Man, and Cyberbetics, vol. 4. pp. 41–46 (1999)
13. Schmidt, C., Kraiss, K.: Establishment of personalized templates for automatic signature verification. Proceedings of the Fourth International Conference on Document Analysis and Recognition, vol. 1, pp. 263–267 (1997)
14. Lee, W.-S., Mohankrishnan, N., Paulik, M.: Improved Segmentation through dyanmic time warping for signature verification

using a neural network classifier. ICIP 98. Proceeding of the 1998 International Conference on Image Processing, vol. 2, pp. 929–933 (1998)
15. Wu, Q,-Z., Jou, I-C Lee, S,-Y.: On-line signature verification using LPC cepstrum and neural networks. IEEE Trans. Syst. Man Cybern., Part B, **27**, pp. 148–153 (1997)
16. Mingming, M., Wijesona, W.: Automatic on-line signature verification based on multiple models., 2000. Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFFr), pp. 30–33 (2000)
17. Han C.-C., et al.: An on-line signature verification system using multi-template matching approaches Proceedings of the IEEE 33rd Annual 1999 International Carnahan - Conference on Security Technology, pp. 477–480 (1999)
18. Yeung, D., et al.: SVC2004: First International Signature Verification Competition In LNCS, Vol. 3072, pp. 16–22, Springer (2004)
19. Lee, L., Berger, T., Aviczer, E.: Reliable online human signature verification systems. IEEE Trans. Syst. Man Cyber. **18**, 643–647 (1996)
20. Justino, E., Bortolozzi, F., Sabourin, R.: The interpersonal and intrapersonal variability influences on off-line signature verification using HMM. Proceedings of the XV Brazilian Symposium on Computer Graphics and Image Processing, pp. 197–202 (2002)

Zero Effort Impostor Attempt

► Zero Effort Forgery

Zone

The print Roman characters usually occupy three zones: the upper, the middle and the bottom zones. For instance, letters “a,” “c,” “e,” “m,” etc., appear in the middle zone; ascendant letters “b,” “d,” “h,” etc., occupy both the upper zone and the middle zone; descendent letters “g,” “p,” etc., occupy both the middle zone and the bottom zone. There are some letters that may cover all the three zones such as “f” and “j.”

► Signature Sample Synthesis

List of Entries

List of Entries

Essays are shown in blue

Abstract Syntax Notation One
ACBio instance
Acceleration
Access Control
Access Control, Logical
Vance Bjorn
Access Control, Physical
Colin Soutar
Accessibility
ACE-V
Action Categorization
Action Understanding
Active (Contour, Shape, Appearance) Models
Acupuncture
AdaBoost
Adapted Fusion
Adaptive Learning
Affective Computing
Albedo
Alignment
Altitude
Ambient Space
American National Standards Institute (ANSI)
Anatomy
Anatomy of Eyes
Kristina Irsch, David L. Guyton
Anatomy of Face
Anne M. Burrows, Jeffrey F. Cohn
Anatomy of Fingerprint
Anatomy of Friction Ridge Skin
R. Austin Hicklin
Anatomy of Hand
Amioy Kumar, Tanvir Singh Mundra, Ajay Kumar
Analysis-by-Synthesis
Analytic Study
And-Or Graph
And–Or Graph Model for Faces
Feng Min, Jinli Suo, Song-Chun Zhu
Annotated Face Model
Anthropometry
Anthroposcopy
Anti-Spoofing
Appearance-Based Gait Analysis
Application Programming Interface (API)
Artifact
Artificial Biometrics
Artificial Digital Biometrics
Artificial Fingerprints
Artificial Image Biometrics
ASN.1
Asset Protection
Association
Attack Trees
Audio–Visual-Dynamic Speaker Recognition
Audio-Visual Fusion
Audio-Visual Speaker Recognition
Audio-Visual Speech Processing
Authentication
Authentics Distribution
Automated Fingerprint Identification System
Automatic Classification of Left/Right Iris Image
Yung-hui Li, Marios Savvides
Average Correlation Energy (ACE)
Azimuth
Background Checks
Peter T. Higgins
Background Subtraction
Back-of-Hand Vascular Pattern Recognition
Back-of-Hand Vascular Recognition
Alex Hwansoo Choi
Barefoot Morphology Comparison
Base Classifier
Baseline Algorithm
Baum-Welch Algorithm
Bayes Decision Theory
Bayes Rule
Bayesian Approach/Likelihood Ratio Approach
Bayesian Hypothesis Test
Behavioral Biometrics
BIAS
Bias-Variance Decomposition
Bi-directional Reflectance Distribution Function (BRDF)
Bifurcation
Binary Hypothesis
Binary Morphology
Binomial Distribution

© 2009 Springer Science+Business Media, LLC

BioAPI
BioAPI Framework
BioAPI Interworking Protocol
Biological Motion
Biometric Algorithms
Yi Chen, Jean-Christophe Fondeur
Biometric and User Data, Binding of
Peng Li, Jie Tian, Xin Yang, Sujing Zhou
Biometric Applications, Overview
David Day
Biometric Capture
Biometric Capture Device
Biometric Characteristic
Biometric Cryptosystem
Biometric Data
Biometric Data Acquisition
Biometric Data Block (BDB)
Biometric Data Capture
Biometric Data Interchange Format
Biometric Data Interchange Format, Standardization
Christoph Busch, Greg Canon
Biometric Data Interchange Record (BDIR)
Biometric Data Interchange Standard
Biometric Decision Time, and the External Operation Time
Biometric Devices
Biometric Encryption
Biometric Engines
Biometric Features
Biometric Fraud Reduction
Biometric Front End
Biometric Fusion
Biometric Fusion Standardization
Biometric Fusion, Rank-Level
Biometric Header, Standards
Biometric Identity Assurance Services
Matthew Swayze
Biometric Information Record
Biometric Interchange Formats
Biometric Interfaces
Catherine J. Tilton
Biometric Key Generation
Biometric Locking
Biometric Match-on-Card, MOC
Biometric Modality
Biometric PAC
Biometric Performance Evaluation Standardization
Biometric Quality Evaluation
Biometric Quality Standards
Biometric Readers
Biometric Recognition
Biometric Reference
Biometric Registration Authority
Biometric Sample
Biometric Sample Acquisition
Dale Setlak
Biometric Sample Quality
Elham Tabassi, Patrick Grother
Biometric Sample Quality, Standardization
Elham Tabassi, Patrick Grother
Biometric Sample Synthesis
Douglas J. Buettner
Biometric Security Measure
Biometric Security, Standardization
Greg Cannon, Philip Statham, Asahiko Yamada
Biometric Security Threat
Biometric Sensing
Biometric Sensor and Device, Overview
Geppy Parziale
Biometric Sensors
Biometric Services
Biometric Specific Threats
Biometric Spoof Prevention
Biometric Strength of Function
Biometric Subsystem Transaction Time
Biometric System
Biometric System Components
Biometric System Design, Overview
Anil K. Jain, Karthik Nandakumar
Biometric Systems, Agent-Based
Farzin Deravi
Biometric Technical Interface, Standardization
John Larmouth
Biometric Technology Test
Biometric Template
Biometric Terminal
Biometric Transaction Time
Biometric Variability
Biometric Verification/Identification/Authentication/Recognition: The Terminology
James L. Wayman
Biometric Vocabulary, Standardization
Rene McIver
Biometric Vulnerabilities, Overview
Andy Adler, Stephanie Schuckers
Biometric Watermarking
Biometrics, Overview
Arun Ross, Anil K. Jain

BIP
BIR
Blind Source Separation
Blood Vessel Wall
Brachycephalic
Branch-and-Bound Search
Breeder Documents
Calibration
Camera
Camera Device
Camera Model
Camera Point of View
Cancelable Biometrics
Andy Adler
Canonical Face Model
CANPASS
Capillary Blood Vessel
Capture Volume
Casts
CBEFF
CBEFF Biometric Data Block (BDB)
CBEFF Biometric Information Records (BIRs)
CBEFF Security Block (SB)
CBEFF Standard Biometric Header (SBH)
CBEFF Wrapper
Central Retinal Artery and Vein
Cepstrum Transform
Chaff Points
Challenge Response
Charge Coupled Device (CCD)
Chrominance
Circular Hough Transform
Circumstantial Identification
Classification
Classifier Cascade
Classifier Combination
Classifier Fusion
Classifier Selection
CLEAR
Client
Closed-Set Identification
CMOS Sensor
Color Constancy
Commensurability
Committee-Based Learning
Common Biometric Exchange Formats Framework Standardization
Fernando L. Podio, Fred Herr
Common Feature Approach
Comparison
Comparison Prints
Complementary Metal Oxide Semiconductor (CMOS)
Compliance
Computational Iris Recognition Systems
Concept Drift
Confidence Interval
Configural Processing
Configuration Issues, System Design
Kai Cao, Jie Tian, Yangyang Zhang, Xin Yang
Conformance Testing
Conformance Testing for Biometric Data Interchange Formats, Standardization of
John W. M. Campbell, Gregory Zektser
Conformity
Contact Microphones
Contact-Based
Contactless
Contextual Biases
Continuous Classification
Contour Detection
Contrast
Convenience Sample
Convergence Feature Extraction
Copula
Core
Correct Index Power
Correct Reject Power
Correlation
Correlation Map
Correspondence
Cost Function
Countermeasures
Counter Sign
Covariate
Covariate Studies
Craniofacial Reconstruction
Craniofacial Superimposition
Credential Hardening
Credentialing System
Credit Check
Crew Designs
Criminal History Check
Criminal Law Enforcement
Criminal Record Search
Cross-Modality Face Biometrics
Cross-Validation
Cryptography

Curse of Dimensionality
Curse of Misalignment
Cursive
Custody Suite
Cut Finger Problem
Dactyloscopist: Fingerprint Examiner
Data Hiding
Data Interchange Format
Data Protection
Database Filtering
Daubert Standard
Daugman Algorithm
Dead Finger Detection
Decision
Decision Criterion Adjustment
Deformable Models
Thomas Albrecht, Marcel Lüthi, Thomas Vetter
Deformation
Delta
Demisyllables
Dental Biometrics
Hong Chen, Anil K. Jain
Dental Identification
Deployment
Depth of Field (DOF)
Dermis
DET Curves
Detector – Extractor
Diffraction Limit
Diffuse Reflection
Digital Watermarking
Digitizer
Digitizing Tablet
Sonia Garcia-Salicetti, Nesma Houmani
Dimensionality Reduction
Diode
Diphones
Disclosure Check
Discriminative Classifier
Dissimilarity
Distortion
Distributed Computing
Distributed Detection
Distributed Inference Making
DNA Analysis
DNA Fingerprinting/DNA Profiling
DNA Profiling
DNA Typing
Dolicocephalic
Double Angle Representation
Double Dipping
Drive-up
Duplicate Detection
Dynamic Programming Comparison Method
Dynamic Time Warping (DTW)
Ear Biometrics
Michał Choraś
Ear Biometrics, 3D
Bir Bhanu, Hui Chen
Ear Recognition
Earmark(s)
Earprints
Earprints, Forensic Evidence of
Christophe Champod
e-Authentication, Remote Access (Partial)
Eigenface
Elastically Adaptive Deformable Model
Electromagnetic Radiation
Electromagnetic Resonance
Electromagnetic Spectrum
Embedded Processor
Embedded Software
Embedded Systems
Naohisa Komatsu, Manabu Nakano
Embedding Space
Empirical Analysis
Empirical Statistical Models
Encoded Finger Data
Encoder
Encoding of Hand Geometry Information
Encryption, Biometric
Ann Cavoukian, Alex Stoianov
Enhancement
Enrollment
Enrollment Time
Enrollment Transaction Duration
Ensemble Learning
Zhi-Hua Zhou
Entropy, Biometric
ePassport
Epidermis
Ergonomic Design for Biometric Systems
Eric P. Kukula, Stephen J. Elliott
Ergonomics
Error Probability Non-Accumulation
Evaluation of Gait Recognition
Sudeep Sarkar, Zongyi Liu

Expected Performance or Utility of Fingerprint Image in an Automated Comparison Environment
Expression
Extended Enterprise
External Identification
External Operation Time
Extra-Class
Eye Centers
Eye Tracking
Face Acquisition
Face Aging
Face Alignment
Leon Gu, Takeo Kanade
Face Alignment Error
Face Biometric
Face Camera
Face Databases and Evaluation
Dmitry O. Gorodnichy
Face Detection
Ming-Hsuan Yang
Face Device
Massimo Tistarelli
Face Identification
Face Image Data Interchange Formats
Face Image Data Interchange Formats, Standardization
Patrick Grother, Elham Tabassi
Face Image Quality Assessment Software
Face Image Synthesis
Face Localization
Face Matching
Face Misalignment Problem
Shiguang Shan, Xilin Chen, Wen Gao
Face Photograph
Face Pose Analysis
Ioannis Patras
Face Pose Estimation
Face Pose Recognition
Face Processing
Face Recognition
Face Recognition From Image Sequences
Face Recognition in Near-Infrared Spectrum
Face Recognition Performance Evaluation
Face Recognition Using Local Features
Face Recognition, 3D-Based
Ioannis A. Kakadiaris, Georgios Passalis, George Toderici, Takis Perakis, Theoharis Theoharis
Face Recognition, Component-Based
Onur C. Hamsici, Aleix M. Martinez
Face Recognition, Geometric vs. Appearance-Based
Lior Wolf
Face Recognition, Near-Infrared
Stan Z. Li, Dong Yi
Face Recognition, Overview
Aleix M. Martinez
Face Recognition, Thermal
George Bebis
Face Recognition, Video-Based
Rama Chellappa, Gaurav Aggarwal
S. Kevin Zhou
Face Reconstruction
Face Registration
Face Sample Quality
Kui Jia, Shaogang Gong
Face Sample Standardization
Face Sample Synthesis
Sami Romdhani Jasenko Zivanov
Face Sample Utility
Face Sketching
Face Tracking
Amit K. Roy-Chowdhury, Yilei Xu
Face Variation
Carlos D. Castillo, David W. Jacobs
Face Verification
Face Warping
Face, Forensic Evidence of
Michael C. Bromby
Facial Action Coding
Facial Changes
Facial Expression Analysis
Facial Expression Recognition
Maja Pantic
Facial Landmarks
Facial Mapping
Facial Motion Estimation
Facial Photograph
Factor Analysis
Failure to Acquire Rate
Failure-to-Enrol Rate
Fake Finger Detection
False Match Rate
False Negative Rate
False Non-Match Rate
False Positive Rate
Feathering
Feature Detection
Feature Extraction

Feature Fusion
Feature Map
Feature Selection
Feature Vector
Features
Features vs. Templates
Fidelity
Field of View (FOV)
Finger Data Interchange Format, Standardization
Raul Sanchez-Reillo, Robert Mueller
Finger Geometry, 3D
Sotiris Malassiotis
Finger Pattern Spectral Data
Finger Vein
Hisao Ogata Mitsutoshi Himaga
Finger Vein Authentication Device
Finger Vein Biometric Algorithm
Mitsutoshi Himaga
Finger Vein Feature Segmentation
Finger Vein Imaging Device
Finger Vein Pattern Imaging
Mitsutoshi Himaga
Finger Vein Reader
Mitsutoshi Himaga
Finger Vein Scanner
Fingermark Identification Procedure
Fingerprint
Fingerprint Analysis
Fingerprint Authentication
Fingerprint Benchmark
Fingerprint Binarization
Fingerprint Biometric
Fingerprint Capture
Fingerprint Characteristics
Fingerprint Classification
Xudong Jiang
Fingerprint Comparing
Fingerprint Compression
Nigel M. Allinson
Fingerprint Contrast Enhancement
Fingerprint Corpora
Fingerprint Data Interchange Format
Fingerprint Databases and Evaluation
Fernando Alonso-Fernandez, Julian Fierrez
Fingerprint Device
Fingerprint Encryption
Fingerprint Fake Detection
Jean-François Mainguet
Fingerprint Features
Josef Bigun
Fingerprint Identification
Fingerprint Image Compression
Fingerprint Image Enhancement
Masanori Hara
Fingerprint Image Quality
Elham Tabassi, Patrick Grother
Fingerprint Indexing
George Bebis
Fingerprint Individuality
Fingerprint Matching, Automatic
Jie Tian, Yangyang Zhang, Kai Cao
Fingerprint Matching, Manual
Herman Bergman, Arie Zeelenberg
Fingerprint Pre-Matching
Fingerprint Quality
Fingerprint Reading
Fingerprint Recognition, Overview
Davide Maltoni
Fingerprint Representation
Fingerprint Retrieval
Fingerprint Sample Synthesis
Raffaele Cappelli
Fingerprint Scan
Fingerprint Sensor
Fingerprint Signatures
Fingerprint Singularity
Fingerprint Skeletonization
Fingerprint Templates
Wei-Yun Yau
Fingerprint Thinning
Fingerprint, Forensic Evidence of
Didier Meuwly
Fingerprint, Palmprint, Handprint and Soleprint Sensor
Geppy Parziale
Fingerprints Hashing
Jean-François Mainguet
First Level Detail
Fisher Criterion
Fixed Pattern Noise
Focal Distance
Focal Length
Footprint Comparison
Footstep Identification
Footstep Recognition
Ruben Vera Rodriguez, Nicholas W. D. Evans, John S. D. Mason

Footstep Verification
Footwear Marks
Footwear Recognition
Maria Pavlou, Nigel M. Allinson
Force Field Feature Extraction
Force Field Transform
Forensic
Forensic Anthropology
Forensic Applications, Overview
Christophe Champod
Forensic Barefoot Comparisons
Brian A. Yamashita, Robert B. Kennedy
Forensic DNA Evidence
T. Hicks, R. Coquoz
Forensic Evaluation of Fingerprints and Fingermarks
Forensic Identification Based on Dental Radiographs
Forensic Science
Forensic Speaker Recognition
Forgery Attempt
Forgery Sign
Forward-Backward Algorithm
Fourier Transform
Fovea
Fragile Bits
Fraud Deterrence
Fraud Mitigation
Fraud Reduction, Applications
Victor Minchih Lee
Fraud Reduction, Overview
Victor Minchih Lee
Freeman Chain Code (FCC)
Function Creep
Fundamental Frequency, Pitch, F0
Fusion Network Topology
Fusion, Biometric
Fusion, Confidence Level
Fusion, Data Level
Fusion, Decision-Level
Lisa Osadciw, Kalyan Veeramachaneni
Fusion, Feature-Level
Arun Ross
Fusion, Image Level
Fusion, Measurement Level
Fusion, Physics-Based
Fusion, Quality-Based
Norman Poh
Fusion, Rank-Level
Ajay Kumar
Fusion, Score-Level
Arun Ross, Karthik Nandakumar
Fusion, Sensor-Level
Afzel Noore, Richa Singh, Mayank Vasta
Fusion, User-Specific
Julian Fierrez, Javier Ortega-Garcia
Fusion, Wavelet-Based
Fuzzy Extractor
Fuzzy Vault
Gabor Jets
Gabor Transform
Gabor Wavelets
Gait
Gait Analysis
Gait Biometrics, Overview
Rama Chellappa, Ashok Veeraraghavan, Narayanan Ramanathan
Gait Models for Biometrics
Gait Recognition
Gait Recognition, Model-Based
Chew-Yean Yam, Mark S. Nixon
Gait Recognition, Motion Analysis for
Ahmed Elgammal
Gait Recognition, Silhouette-Based
Jeffrey E. Boyd, James J. Little
Gait, Forensic Evidence of
Niels Lynnerup, Peter K. Larsen
Gallery and Probe
Gaussian Mixture Density
Gaussian Mixture Models
Douglas Reynolds
GC
Gelatin Pad
General Model
Generalization
Generalization Error
Generative Classifier
Genetic Identification
Genuine Matching
Genuine Sign
Genuine/Impostor Attempt
Geodesic
Geometry Image
Global Fusion
Global Thresholding Techniques
Glottal Excitation
Glyph
GMM
Graph Matching

Graphic Tablet
Graphical User Interface
Graphometric Features
Gray Scale
GRF (Ground Reaction Force)
Ground-Truth
Gummy Bear Finger
Haar-Like Features
Habituated Subject
Habituation
Halo Effect
Hamming Distance
Hand Biometrics
Hand Biometrics, 3D
Hand Contour
Hand Data Interchange Format, Standardization
Raul Sanchez-Reillo, Samir Tamer
Hand Databases and Evaluation
Guangming Lu
Hand Geometry
Raul Sanchez-Reillo
Hand Geometry View Record – HGVR
Hand Physiology
Hand Shape
Nicolae Duta
Hand Shape Biometrics
Hand Silhouette Data
Hand Structure
Hand Vascular Recognition
Hand Veins
Graham Leedham
Hand Vein Identification
Hand Vein Verification
Hand-Geometry Device
Vitomir Štruc, Nikola Pavešić
Hand-Geometry Reader
Hand-Geometry Scanner
Hand-Held Devices
Handprint
Handprint Sensor
Hand-Scanning Device
Handwriting Sample Synthesis
Handwriting Synthesis
Handwritten Signature Recognition
Head Pose Analysis
Head Yaw/Tilt/Roll
Headspace
Helper Data
Heterogeneous
Heterogeneous Face Biometrics
Stan Z. Li
Heterogenous Face Image Matching
Hidden Markov Models
Javier Hernando
Hill-Climbing Attack
Histogram Equalization
HMM
Human Computing
Human Dental Atlas
Human Detection and Tracking
James W. Davis, Vinay Sharma, Ambrish Tyagi, Mark Keck
Human Factors
Human Movement, Psychology
Human-Biometric Sensor Interaction (HBSI)
Human–Computer Interaction (HCI) and User Interfaces
Human-Interpretable Fingerprint Classes
Hypothesis Test
ICP Algorithm
ID Photograph
Identification
Identity Level in the Speech Signal
Identity Theft Reduction
Illumination
Illumination Compensation
Xudong Xie, Kin-Man Lam, Qionghai Dai
Illumination Normalization
Image Acquisition
Image Capture
Image Classification
Image Enhancement
Image Formation
Xiaoming Peng
Image Formation Process
Image Morphology
Image Pattern Classification
Image Pattern Recognition
Jian Yang, Jingyu Yang
Image Recognition
Image Regeneration from Templates
Image Resolution
Image Segmentation
Image Warping
Imaging Spectroscopy
Imaging Volume
Implementation Under Test (IUT)
Impostor

Imposter Distribution
Impostor Match
Imprecise Localization
Incremental Learning
Xin Geng, Kate Smith-Miles
Independent Component Analysis
Seungjin Choi
Independent Factor Analysis
Indexing
Individuality of Biometric Traits
Individuality of Fingerprints
Sarat C. Dass, S. Pankanti, S. Prabhakar, Y. Zhu
Individualization
Influencing Factors
Influential Factors to Performance
Kaoru Uchida
Information Content of Iris Images
Information Fusion
Intelligent Agents
Interaction
Interactive Voice Response (IVR)
Interest Point, Region, Local Feature
Interest Points
Interface
Intermediate Biometrics
Internal Identification
International Association for Identification
Interoperability
Interoperable Performance
Patrick Grother
Intraclass
Intricated
Intricated Biometrics
Intrinsic Dimensionality of a Manifold
Intrinsic Direction of Fingerprint
Intrinsic Distance
Intrinsic Failure
Invariant–Covariant
Iris
Iris Acquisition Device
Ryan Rakvic, Randy Broussard, Lauren Kennell, Robert Ives, Robert Bell
Iris at a Glance
Iris Biometric
Iris Camera
Iris Capture
Iris Data Interchange Standards
Iris Databases
Damon L. Woodard, Karl Ricanek
Iris Device
James R. Matey
Iris Digital Watermarking
Nick Bartlow, Nathan Kalka, Bojan Cukic, Arun Ross
Iris Encoding and Recognition using Gabor Wavelets
John Daugman
Iris Image Capture Device
Iris Image Data Interchange Formats, Standardization
James L. Cambier
Iris Image Enhancement by Super-Resolution Method
Iris Image Quality
Natalia A. Schmid
Iris Interchange Format Standards
Iris Localization
Iris on the Move™
James R. Matey
Iris Quality Metrics
Iris Reader
Iris Recognition, Overview
Yung-Hui Li, Marios Savvides
Iris Recognition Algorithms
Iris Recognition at Airports and Border-Crossings
John Daugman
Iris Recognition Immigration System (IRIS)
Iris Recognition Operational Range
Iris Recognition Performance Under Extreme Image Compression
John Daugman, Cathryn Downing
Iris Recognition Systems
Iris Recognition Using Correlation Filters
Yung-hui Li, Marios Savvides, Jason Thornton, B. V. K. Vijaya Kumar
Iris Recognition with Deformation and Occlusion Estimation
Iris Retina Biometric Fusion
Iris Sample Synthesis
Natalia A. Schmid
Iris Scan
Iris Scanner
Iris Segmentation
Iris Segmentation Using Active Contours
Sung W. Park, Marios Savvides
Iris Segmentation Using Snakes
Iris Standards Evolution
Iris Standards Progression
Dominique Harrington, Ryan Triplett

Iris Super-Resolution
Yung-Hui Li, Marios Savvides
Iris Template Extraction Via Bit Inconsistency and GRIT
Gerry Vernon Dozier, Marios Savvides, Kelvin Bryant, Taihei Munemoto, Karl Ricanek, Jr., Damon Woodard
Iris Template Protection
Patrizio Campisi, Emanuele Maiorana, Alessandro Neri
Iris Template Security
Iris2pi
IrisCode
ISO
JPEG and JPEG2000 Image Compression
Kernel
Key Binding
Keypoints
Keystroke Dynamics
Keystroke Pattern Classification
Keystroke Recognition
Nick Bartlow
Kinematic Body Model
Kinematics
Knowledge-based Gait Recognition
Known Traveler
L2 norm
Lambertian Law
Lambertian Surface
Large Scale Biometric Database
Large Scale Biometric System Design
Large Scale System Design
Chin-Hung Teng, Wen-Hsing Hsu
Large-Scale Evaluation
Latent Fingerprint
Latent Fingerprint Experts
Thomas A. Busey, Bethany L. Schneider
Latex Finger
Law Enforcement
Ken Moses
Law Enforcement Agency
LBP (Local Binary Pattern)
LCN DNA/Low Template Level
LDA (Linear Discriminant Analysis)
LED (Light Emission Diode)
Lighting Compensation
Lighting Model
Likelihood Ratio Test
Limbus
Linear Dimension Reduction
Wei-Shi Zheng, J. H. Lai, Pong C. Yuen
Linear Feature Extraction
Linearly Symmetric Image
Lip Movement Recognition
Petar S. Aleksic
Lip-Radiation Effect
Liquid Crystal Displays (LCD)
Liveness Detection
Liveness Assurance in Face Authentication
Michael Wagner, Girija Chetty
Liveness Assurance in Voice Authentication
Michael Wagner
Liveness Detection
Liveness Detection
Liveness Detection: Fingerprint
Stephanie A. C. Schuckers
Liveness Detection: Iris
Bori Toth
Live-Scan Furrow Device
Live-Scan Sensor
Local Adaptive Thresholding
Local Descriptors
Local Fusion
Local Image Features
Krystian Mikolajczyk, Tinne Tuytelaars
Local Image Filters
Abdenour Hadid, Matti Pietikäinen
Local Surface Patch
Localization
Localization Inaccuracy
Logical Access Control, Client-Based
Logical Access Control, Client-Server-Based
Logico-Linear Operator
Logon, Password Management
Luminance
Machine-Generated Fingerprint Classes
Machine-Learning
Magnification
Mahalanobis Distance
Malicious-code-free Operating System
Manifold
Manifold Embedding
Manifold Learning
Philippos Mordohai, Gérard Medioni
Manual Annotation
Margin Classifier

Markerless 3D Human Motion Capture from Images
P. Fua
Match Score Fusion
Matcher
Matching
Matching Score
Match-On-Card
Maximum A Posteriori (MAP)
Maximum A-Posteriori Estimation
Maximum Likelihood Estimation
Maximum Margin Classifier
Maximum Permissible Exposure (MPE)
Mesocephalic
Metatarsal Ridge
Microphone
Microphone Arrays
Minimal Constraint Iris Recognition
Minutia
Minutia Direction
Mislabeled Iris Data Correction
Mitochondrial DNA
Mixture Mode
Model-Based Biometrics
Monitoring
Monomodal/Multimodal Database
Morphable Models
Mosaicing
Motion Capture
Motion Estimation
Motion Model
Motion Recovery, 3D
Moving Light Display
MS
Multi-Algorithm Systems
Multi-Instance Systems
Multi-Modal Samples
Multi-Sample Systems
Multi-Sensor Systems
Multi-Unit Systems
MultiBand Biometrics
Multibiometric Fusion, Standardization
MultiBiometric Systems
Multibiometrics
Arun Ross
Multibiometrics and Data Fusion, Standardization
Jung Soh, Farzin Deravi, Alessandro Triglia, Alex Bazin
Multifactor
Multimodal
Multimodal Fusion
Multimodal Jump Kits
Multiple Classifier Fusion
Multiple Classifier Systems
Fabio Roli
Multiple Classifiers
Multiple Expert Systems
Multiple Experts
Josef Bigun
Multiple View Geometry
Multispectral and Hyperspectral Biometrics
Besma Roui-Abidi, Mongi Abidi
Multistage Matching
Mutual Authentication
Mutual information
NAP-SVM
National Institute for Standards and Technology
Natural Gradient
Near Field Communication
Near Infrared (NIR)
Near-infrared Image Based Face Recognition
NEXUS
NIST SREs (Speaker Recognition Evaluations)
Noisy Iris Challenge Evaluation – Part I (NICE.I)
Nominal Identity
Non-ideal Iris
Non-linear Dimension Reduction Methods
Non-linear Techniques for Dimension Reduction
Jian Yang, Zhong Jin, Jingyu Yang
Normalised Hamming Distance
Nuisance Attribute Projection
Numerical Standard
Object Recognition
Observations from Speech
Ocular Biometrics
Odor Biometrics
Adee A. Schoon, Allison M. Curran, Kenneth G. Furton
Off-Angle or Nonorthogonal Segmentation
On-Card Matching
Chen Tai Pang, Yau Wei Yun, Xudong Jiang
One-to-Many Identification
One-to-One Verification
Online Learning
Open-Set Identification

Operational Tests
Operational Times
Stephen J. Elliott, Eric P. Kukula, Richard T. Lazarick
Optical Flow
Optical Target
Optimal Hyperplane
Optimization
Ordinal Measure
Orthographic Scanning
Osmology
Output Noise Variance (ONV)
Outsole Pattern Matching
Overfitting
Paleoanthropology
Palm Dorsal Vein
Palm Segment
Palm Vein
Masaki Watanabe
Palm Vein Authentication
Palm Vein Authentication Sensor
Palm Vein Image Sensor
Masaki Watanabe
Palm Vein Recognition
Palm Vein Scanner
Palmprint
Palmprint, 3D
David Zhang, Vivek Kanhangad
Palmprint Anatomy
Palmprint Authentication, 3D
Palmprint Characteristics
Palmprint Database
Palmprint Device
Palmprint Features
David Zhang, Laura L. Liu
Palmprint Matching
Andrew Beng Jin Teoh
Palmprint Recognition, 3D
Palmprint Representation
Palmprint Sensor
Parallel Fusion Network
Parametric Models
Parametric-Based Biometrics
Part-Based Face Recognition
Partial Occlusion
Passive Biometrics
Patron Format Specification
Pattern Recognition
PCA (Principal Component Analysis)
Pedestrian Detection
Peg
Pen Altitude
Pen Azimuth
Pen Inclination (Pen Tilt)
Pen Pressure
Pen Tablet
Penetration Rate
Pen-tip Position (Pen Coordinates)
Perceptual Expertise
Performance Bias in Synthesized Biometric Data
Performance Evaluation Measures
Performance Evaluation, Overview
Shiguang Shan, Xilin Chen, Wen Gao
Performance Measures
Jihyeon Jang, Hale Kim
Performance Metrics
Performance of Biometric Quality Measures
Performance Testing
Performance Testing Methodology Standardization
Michael Thieme
Perpetrator Identification
Personal Data
Personal Information Search
Person-Independent Model
Phase
Phoneme
Photogrammetry
Photography for Face Image Data
Ted Tomonaga
Photography Guidelines
Photometric Guidelines
Photon
Physical Analogies for Ear Recognition
David J. Hurley, Mark S. Nixon
Physical and Logical Access Control convergence
Physics-Based Biometrics
Physics-Based Models
Piezoelectric
PIN Replacement
Pitch
Pixel
Platen
Point-Light Display
Polar
Polarized
Police Law Enforcement

Portal
Pose
Pose and Motion Models
Poststratification
Potential Energy Transform
Preemployment Screening
Pre-Processing
Pretty Good Privacy (PGP)
Primary Biometric Identifier
Principal Component Analysis
Principal Curvatures
Principal Lines
Privacy
Privacy Issues
Terence M. Sim
PrivateID™
Privium
Probabilty Density Function (PDF)
Process Artifacts
Procrustes Shape Distance
Proposal Descriptive and Decision Making Model
Prosody
Protocol
Pseudo-Random Number Generator
Psychology of Gait and Action Recognition
Frank E. Pollick
Punch-in, Clock-in, Punch-out, Clock-out, Punch
Punctum Lacrimale
Pupil
Pupil Phase Engineered Iris Biometrics
Pupil Phase Engineering
Purkinje Images
Quadrant
Quality-dependent Fusion
Quantum Efficiency (QE)
Radiometric Calibration
RAIC
Random Forgery
Range Scans
RASTA-Filtering
Raw Finger Vein Image
Read Noise
Real-Time 3D Surface Digitization
Recognition at a Distance
Rectilinear
Reference Set
Reflection
Reflection-Based Touchless Finger Imaging (RTFI)
Region-of-Interest (ROI) Encoding
Registered Traveler
Catherine J. Tilton
Remote Authentication
Judith Markowitz
Remote Monitoring (Partial)
Remote Verification
Rendering
Replay Attack
Residence Time
Resolution
Response Time
Retina
Retina Recognition
Yoichi Seto
Retinal Angiogenesis
Retinal Blood Vessels
Retinal Scan
Reverse Engineering
Revocable Biometrics
Ridge Enhancement
Ridge Extraction
Ridge Flow
Robustness Test
ROC Curve
Rolled-Equivalent Fingerprint
Rolls Capture Device
Rotation Angle
Sample Quality
Sample Size
Sampling Frequency
Scalability
Scenario Tests
Scene Marks
Scent Identification Line-Ups
Score Fusion
Score Fusion and Decision Fusion
Score Normalization
Score Normalization Rules in Iris Recognition
John Daugman
Sealed Local Biometric Identity Verification Systems
Second Level Detail
Secure Biometric Token Operating System
Secure Biometrics
Secure Element
Secure Sketch
Security and Liveness, Overview
Andy Adler, Stephanie Schuckers

Security Block
Security Issues, System Design
Karthik Nandakumar
Security Threat Assessment
Segmentation of Iris Images from Noncooperative Subjects
Segmentation of Off-Axis Iris Images
Lauren R. Kennell, Ryan N. Rakvic, Randy P. Broussard
Semi-Supervised
Sensitivity Analysis in Biometric Systems
Service-Oriented Architecture
Service Provider
Session and Channel Variabilities
Session Effects on Speaker Modeling
Driss Matrouf, Jean-François Bonastre
SFinGe
Raffaele Cappelli
Shape
Shape Index
Shape Model
Shape vs. Texture
Shielding Functions
Shoeprint Matching
Shot Noise
Side-Channel Attacks
Signal to Noise Ratio
Signature Benchmark
Signature Characteristics
Signature Corporate
Signature Databases and Evaluation
Marcos Martinez-Diaz, Julian Fierrez
Signature Dataset
Signature Features
Marcos Martinez-Diaz, Julian Fierrez, Seiichiro Hangai
Signature Matching
Marcos Martinez-Diaz, Julian Fierrez, Seiichiro Hangai
Signature Recognition
Olaf Henniger, Daigo Muramatsu, Takashi Matsumoto, Isao Yoshimura, Mitsu Yoshimura
Signature Sample Synthesis
Liang Wan, Zhouchen Lin
Signature Similarity Computation
Signature Synthesis
Signature/Sign Recognition
Silhouette
Silhouette Analysis for Gait Recognition
Similarity Metric
Simplifying Passenger Travel Program
Simultaneous Capture of Iris and Retina for Recognition
David Usher, Yasunari Tosa, Marc Friedman
Skilled Forgery
Skin Classification
Skin Color Detection
Skin Detection
Ahmed Elgammal, Crystal Muang, Dunxu Hu
Skin Print
Skin Spectroscopy
Dong Yi, Weilong Yang, Stan Z. Li
Skin Texture
Xiangxin Zhu, Zhen Lei, Stan Z. Li
Skull, Forensic Evidence of
Mingquan Zhou
Skull-Photo Superimposition
Slap Or Four-Four-Two device
Smart Cameras
Smart Card
Soft Biometrics
Karthik Nandakumar, Anil K. Jain
Soleprint Device
Soleprint Sensor
Sound
Sound Generation
Sources of Evidence
Sources of Information in Biometric Fusion
Arun Ross
Speaker Authentication (Partial Synonym)
Speaker Biometrics
Speaker Change Detection
Speaker Classification
Speaker Clustering
Speaker Databases and Evaluation
Alvin F. Martin
Speaker Detection
Speaker Diarization
Speaker Features
Daniel Ramos, Javier Gonzalez-Dominguez, Doroteo T. Toledano, Joaquin Gonzalez-Rodríguez
Speaker Identification and Verification, SIV
Speaker Indexing
Speaker Matching
Jean-François Bonastre, Driss Matrouf
Speaker Model

Speaker Parameters
Speaker Recognition Engine
Speaker Recognition, One to One
Speaker Recognition, Overview
Jean Hennebert
Speaker Recognition, Standardization
Judith Markowitz
Speaker Segmentation
Laura Docio-Fernandez, Carmen Garcia-Mateo
Speaker Separation
Speaker Tracking
Speaker Verification
Spectral Analysis of Skin
Specular Reflection
Specularity
Speech Analysis
Doroteo T. Toledano, Daniel Ramos, Javier Gonzalez-Dominguez, Joaquín González-Rodríguez
Speech Input Device
Speech Parametrization
Speech Processing
Speech Production
Laura Docio-Fernandez, Carmen Garcia-Mateo
Speech Recognition
Speech Spectral Envelope
Speech Synthesis
Speech System
Speed
SPME
Spoofing
Spoofing Countermeasures
Spoof-resistance
Stand Off
Statistical Models
Statistical Signal Processing
Steganography
Strain Gauge
Stream of Speech
Strength of Voice Evidence
Stroke
Structural and Functional Anatomy
Structural Model
Structural Risk
Structure Tensor Field
Structure-from-Motion
Subject Interaction Time
Super-Resolution
Super-Resolution for Iris
Supervised
Supervised Learning
Jong Kyoung Kim, Kye-Hyeon Kim, Seungjin Choi
Supervisor
Supervisor Opinion
Support Vector Machine
Mathias M. Adankon, Mohamed Cheriet
Surface Curvature
Surface Matching
Surveillance
Rama Chellappa, Aswin C. Sankaranarayanan
SVM
SVM Supervector
Sweep Sensor
Synthesis Attack
Synthetic Biometrics
Synthetic Fingerprint Generation
Synthetic Fingerprints
Synthetic Iris Images
Synthetic Voice Creation
System-on-card
Tablet
Tamper-Proof
Tamper-proof Operating System
Raul Sanchez-Reillo
Target Detection
Target Population
Target-Dependent Fusion
Technology Tests
Template
Template Distortion
Template Protection
Template Reconstruction
Template Security
Andy Adler, Raffaele Cappelli
Temporal Characterization of Faces
Temporal Domain
Tenprint Capture
Tensor
Test Sample and Size
Michael E. Schuckers
Text-Dependent
Text-Independent
Text-Prompted
Text-to-Speech (TTS)
Text-to-Speech (TTS) Synthesis
TFIR

Thermal Biometrics
Thermogram
Third Level Detail
Threshold Limit Value (TLV)
Thresholding
Throughput
Time and Attendance
Samir Tamer, Stephen J. Elliott
Time and Attendance Terminal
Time Clock
Time Series
Tongue-Print Recognition
Tooth Biometrics
Top and Secondary Choices
Total Transaction Time
Touch Tablet
Touch-Screen
Tracing
Training
Training Data, Sufficiency
Training Signature
Transfer Learning
Transformation
Transmission-Based Touchless Finger Imaging (TTFI)
Transportable Asset Protection
Anthony P. Russo
Tread Pattern
Trusted Biometric System
Trusted Traveler
UBM
ULW
Unauthorized Data Collection
Unauthorized Data Disclosure
Unification Framework
Universal Background Models
Douglas Reynolds
Universal Latent Workstation
Tom Hopper
Unnecessary Data Collection
Unsupervised
Unsupervised Rank Level Fusion
Unvoiced Sounds
Usability
User Acceptance
Marek Rejman-Greene
User Interfaces
User Interface, System Design
Jianjie Li, Xin Yang, Xunqiang Tao, Jie Tian
User-Centered Design
User-dependent Fusion
Utility
Utterance
Vascular Biometrics
Vascular Image Data Format, Standardization
Alex H. Choi, Jonathan R. Agre
Vascular Network Pattern
Vascular Recognition
Vector Quantization
Vein
Vein Biometrics
Vein Recognition
Velocity (Speed)
Verification
Vetting
Video Camera
Video Surveillance
Video-based Face Recognition
Video-based Motion Capture
Visible Spectrum
Visual Memory
Visual Sensor
Visual-dynamic Speaker Recognition
Vitality
Viterbi Algorithm
VOCs (Volatile Organic Compounds)
Voice Authentication
Voice Biometric
Voice Biometric Engine
Voice Device
Doroteo T. Toledano,
Joaquin Gonzalez-Rodriguez,
Javier Ortega-Garcia
Voice Evidence
Voice Recognition
Voice Sample Synthesis
Juergen Schroeter
Voice Verification
Voice, Forensic Evidence of
Andrzej Drygajlo
Voiced Sounds
Voiceprint
Volunteer Crew
Walk-through
Walk-up
Watermarking, Biometric
Watermarking, Digital

Wavefront Coded Iris Biometric Systems
V. Paúl Pauca, Kelly Smith Faddis
Arun Ross, Joseph van der Gracht,
Todd C. Torgersen
Wavefront Coding
Wavefront Coding for Enhancing the Imaging Volume in Iris Recognition
Paulo E. X. Silveira, Lu Gao, Ramkumar Narayanswamy
Wavelength
Wavelet Transform
Web Services
Work-sharing On-card Matching
World Model
XML Schema Definition
Zero Effort Forgery
Stephen J. Elliott
Zero Effort Impostor Attempt
Zone